A multi-asset investment and consumption problem with transaction costs

David Hobson[1] · Alex S.L. Tse[2] · Yeqi Zhu[3]

Abstract In this article, we study a multi-asset version of the Merton investment and consumption problem with CRRA utility and proportional transaction costs. We specialise to a case where transaction costs are zero except for sales and purchases of a single asset which we call the illiquid asset. We show that the underlying HJB equation can be transformed into a boundary value problem for a first order differential equation. Important properties of the multi-asset problem (including when the problem is well-posed, ill-posed, or well-posed for some values of transaction costs only) can be inferred from the behaviours of a quadratic function of a single variable and another algebraic function.

Keywords Portfolio choice · Transaction costs · Multiple assets · Hamilton–Jacobi–Bellman equation · Free boundary value problem

Mathematics Subject Classification (2010) 91G10 · 93E20

JEL Classification G11

✉ D. Hobson
D.Hobson@warwick.ac.uk

A.S.L. Tse
A.Tse@imperial.ac.uk

Y. Zhu
Yeqi.Zhu@credit-suisse.com

[1] Department of Statistics, University of Warwick, Coventry, CV4 7AL, UK

[2] Department of Mathematics, Imperial College London, London, SW7 2AZ, UK

[3] Credit Suisse, London, UK

1 Introduction

In one of his seminal works, Merton [20] studies an optimal investment/consumption problem faced by a risk-averse agent over an infinite horizon. In an economy in which the single risky asset follows an exponential Brownian motion and the agent has constant relative risk aversion, the optimal strategy is to consume at a rate which is proportional to wealth, and to invest a constant fraction of wealth in the risky asset. The result generalises easily to multiple risky assets.

Constantinides and Magill [19] were the first to incorporate proportional transaction costs into the Merton model. They conjectured that the agent should trade minimally to keep the fraction of wealth invested in the risky asset within an interval. Subsequently, Davis and Norman [9] gave a precise characterisation of the trading strategy in terms of local times and formally proved its optimality by a Hamilton–Jacobi–Bellman (HJB) equation verification argument. Shreve and Soner [23] reproved the results of [9] using viscosity solutions and gave several extensions. These approaches remain the main methods for solving portfolio optimisation problems with transaction costs, although recently a different technique based on shadow prices has been proposed; see Guasoni and Muhle-Karbe [12] for a user's guide. Kallsen and Muhle-Karbe [17], Choi et al. [6] and Herczegh and Prokaj [13] use the dual approach to characterise the solution to the problem with transaction costs and one risky asset.

The results in Davis and Norman [9] are limited to a single risky asset, and it is of great interest to understand how they generalise to multiple risky assets where the related literature is limited. On the computational side in a multi-asset setting, Muthuraman and Kumar [21] use a process of policy improvement to construct a numerical solution for the value function and the associated no-transaction region, Collings and Haussmann [7] derive a numerical solution via a Markov chain approximation for which they prove convergence, and Dai and Zhong [8] use a penalty method to obtain numerical solutions. On the theoretical front, Akian et al. [1] show that the value function is the unique viscosity solution of the HJB equation (and provide some numerical results in the two-asset case), and Chen and Dai [4] identify the shape of the no-transaction region in the two-asset case. Explicit solutions of the general problem remain very rare.

One situation when an explicit solution is possible is the rather special case of uncorrelated risky assets and an agent with constant absolute risk aversion. In that case, the problem decouples into a family of optimisation problems, one for each risky asset; see Liu [18]. Another setting for which some progress has been made is the problem with small transaction costs; see Whalley and Wilmott [26], Janeček and Shreve [16], Bichuch and Shreve [3], Soner and Touzi [24], and, for a recent analysis in the multi-asset case, Possamaï et al. [22]. These papers use an expansion method to provide asymptotic formulae for the optimal strategy, value function and no-transaction region.

Our focus is on optimal investment/consumption problems, but there is a parallel literature on optimal investment problems involving maximising expected utility at a distant terminal horizon; see for example Dumas and Luciano [10] for an explicit solution in the one-asset case, and Bichuch and Guasoni [2] for recent work in a setting similar to ours with liquid and illiquid assets.

In this paper, we consider the problem with a risk-free bond and two risky assets.[1] Transactions in the first risky asset (which we call the liquid risky asset) are costless, but transactions in the second risky asset, which we term the illiquid asset, incur proportional costs. This is also the setting of a recent paper by Choi [5]. The main difference between this paper and Choi [5] is that we analyse the HJB equation, whereas Choi takes the dual approach and studies shadow prices.

This paper is an extension of Hobson et al. [14] which considers a similar problem with a bond and an illiquid asset, but with no other risky assets.[2] Many of the techniques of [14] carry over to the wider setting of the present paper. (Similarly, the paper of Choi [5] extends the work of Choi et al. [6] to include a risky liquid asset.) However, since there are fewer parameters when the financial market includes just one risky asset, the problem in [14] is significantly simpler and much more amenable to a comparative statics analysis. In contrast, the present paper treats the multi-asset problem which has proved so difficult to analyse, albeit in a rather special case. The multi-asset setting brings new challenges and complicates the analysis.

Our first achievement is to show that the problem of finding the free boundaries and the value function can be reduced to the study of a family of solutions to a boundary crossing problem for a first order ordinary differential equation parameterised by its initial value; see Problem 3.2 below. Moreover, the existence or otherwise of a solution to this boundary crossing problem can be reduced to a study of the properties of a quadratic function and a second algebraic function. Given the results of Hobson et al. [14] or Choi et al. [6], this is perhaps not a surprise. Nonetheless, the reduction of the problem to a first order equation is key to all subsequent developments.

Our second main achievement is to give necessary and sufficient conditions on the parameters for the problem to be well-posed (Theorem 4.1), and in those cases to give an expression for the value function (Theorem 4.3). These results extend Choi et al. [6] and Hobson et al. [14] to the case of multiple risky assets. We find that the problem may be well-posed, or it may be ill-posed, or (if $R < 1$) it may be well-posed for large transaction costs and ill-posed for small transaction costs, or (if $R > 1$) it may be well-posed for small transaction costs and ill-posed for large transaction costs.

Our third achievement is to make definitive statements about the comparative statics for the problem. We focus on the boundaries of the no-transaction wedge and the certainty equivalent value of the holdings in the illiquid asset. Among other results, we prove (see Theorem 6.2 and Corollary 6.3 for precise statements) that as the return on the illiquid asset improves, the agent aims to keep a larger fraction of his total wealth in the illiquid asset, in the sense that the critical ratios at which sales and purchases take place are increasing in the return. Conversely, as the agent becomes more impatient, the agent keeps a smaller fraction of wealth in the illiquid asset. Further,

---

[1]More generally, we may have several risky assets on which no transaction costs are payable, and a single illiquid asset. This general case can be reduced to the case with two risky assets, one liquid and one illiquid; see Remark 2.5.

[2]This paper can also be viewed as a development of the results of Hobson and Zhu [15]. The model in [15] includes both a liquid risky asset and an illiquid asset, but assumes that the transaction cost on purchases of the illiquid asset is infinite. This case might be called the "perfectly illiquid" case: the illiquid asset can be sold, but not bought, and the problem is an optimal liquidation problem. This paper extends Hobson and Zhu [15] to allow finite transaction costs and purchases of the illiquid asset.

we prove (Theorem 6.4 and Corollary 6.5) that as the return on the illiquid asset improves, or as the agent becomes less impatient, the certainty equivalent value of the holdings in the illiquid asset increases.

The remainder of the paper is structured as follows. In the next section, we formulate the problem. In Sect. 3, we give heuristics showing how the underlying HJB equation can be converted to a free boundary value problem involving a first order differential equation. Then we can state our main results on the existence of a solution in Sect. 4. In Sect. 5, we discuss the various cases which arise. In Sect. 6, we discuss the comparative statics of the problem, before Sect. 7 concludes. Materials on the solution of the free boundary value problem, the verification argument for the HJB equation, and other lemmas on the analysis of solutions of the differential equations are relegated to the appendices.

2 The problem

The economy consists of one money market instrument paying constant interest rate r and two risky assets, one of which is liquidly traded while the other is illiquid. There are no transaction costs associated with trading in the liquid asset. Meanwhile, trading in the illiquid asset incurs a proportional transaction cost $\lambda \in [0, \infty)$ on purchases and $\gamma \in [0, 1)$ on sales, where not both λ and γ are zero. Define[3] $\xi := \frac{1+\lambda}{1-\gamma} - 1 = \frac{\lambda+\gamma}{1-\gamma} > 0$ to be the round-trip transaction cost. Let $(S, Y) = (S_t, Y_t)_{t \geqslant 0}$ be the price processes of the liquid and illiquid assets, respectively. The price dynamics are given by

$$(S_t, Y_t) = \left(S_0 \exp\left(\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma B_t \right), Y_0 \exp\left(\left(\alpha - \frac{\eta^2}{2}\right)t + \eta W_t \right) \right),$$

where (B, W) is a pair of Brownian motions with correlation coefficient given by $\rho \in (-1, 1)$. Write $\beta := (\mu - r)/\sigma$ and $\nu := (\alpha - r)/\eta$ for the Sharpe ratio of the liquid and illiquid asset, respectively.

Let Θ_t be the number of units of the illiquid asset held by an agent at time t. Then $\Theta_t = \Theta_0 + \Phi_t - \Psi_t$, where $\Phi = (\Phi_t)_{t \geqslant 0}$ and $\Psi = (\Psi_t)_{t \geqslant 0}$ are both increasing, right-continuous and nonnegative processes representing the cumulative units of purchases and sales respectively of the illiquid asset. Let $C = (C_t)_{t \geqslant 0}$ be the nonnegative consumption rate process of the agent and $\Pi = (\Pi_t)_{t \geqslant 0}$ the cash value of holdings in the risky liquid asset. We assume Φ, Ψ, C and Π are progressively measurable. If $X = (X_t)_{t \geqslant 0}$ is the total value of the liquid instruments (cash and the liquid risky asset), then, assuming transaction costs are paid in cash and consumption is from the cash account,

$$dX_t = \left((\mu - r)\Pi_t + rX_t - C_t \right) dt - Y_t(1 + \lambda) \, d\Phi_t + Y_t(1 - \gamma) d\Psi_t + \sigma \Pi_t dB_t.$$

We say that a portfolio (X, Θ) is *solvent* at time t if its instantaneous liquidation value is nonnegative, that is, $X_t + \Theta_t^+ Y_t(1 - \gamma) - \Theta_t^- Y_t(1 + \lambda) \geqslant 0$. A consump-

[3]An agent with cash wealth $1 + \xi$, who uses this cash to buy the illiquid asset before selling it again immediately, ends with unit cash amount. Alternatively we may think of $1 + \xi$ as the ratio of the ask-price to the bid-price.

tion/investment strategy (C, Π, Θ) is said to be *admissible* if the resulting portfolio is solvent at the current time and at all the future time points. Write $\mathcal{A}(t, x, y, \theta)$ for the set of admissible strategies with initial time-t value $X_{t-} = x$, $Y_t = y$, $\Theta_{t-} = \theta$.

We assume the agent has a CRRA utility function with risk aversion parameter $R \in (0, \infty) \setminus \{1\}$ and subjective discount rate δ. The underlying problem of this paper is the following *investment/consumption problem*.

Problem 2.1 Find an admissible (C, Π, Θ) to maximise the agent's expected lifetime discounted utility from consumption, i.e., find

$$V(x, y, \theta) = \sup_{(C, \Pi, \Theta) \in \mathcal{A}(0, x, y, \theta)} \mathbb{E}\left[\int_0^\infty e^{-\delta s} \frac{C_s^{1-R}}{1-R} \, ds\right]. \tag{2.1}$$

We call $X_t + \Theta_t Y_t$ the *paper wealth* of the agent. In our parametrisation, a key quantity will be $P_t := \frac{\Theta_t Y_t}{X_t + \Theta_t Y_t}$, the proportion of paper wealth invested in the illiquid asset. Building on the intuition developed by Constantinides and Magill [19] and Davis and Norman [9], we expect that the optimal strategy of the agent is to trade the illiquid asset only when P_t falls outside a certain interval $[p_*, p^*]$ to be identified. Note that due to the solvency restriction, we must have $-\frac{1}{\lambda} \leqslant P_t \leqslant \frac{1}{\gamma}$, and hence the no-transaction wedge must satisfy $[p_*, p^*] \subseteq [-\frac{1}{\lambda}, \frac{1}{\gamma}]$.

Define the auxiliary parameters b_1, b_2, b_3 and b_4 as

$$b_1 = \frac{2(\delta - r(1-R) - \frac{\beta^2(1-R)}{2R})}{\eta^2(1-\rho^2)}, \qquad b_2 = \frac{\beta^2 - 2R\eta\rho\beta + \eta^2 R^2}{\eta^2 R^2(1-\rho^2)},$$

$$b_3 = \frac{2(\nu - \beta\rho)}{\eta(1-\rho^2)}, \qquad b_4 = \frac{2}{\eta^2(1-\rho^2)}.$$

It turns out that the optimal investment and consumption problem depends on the original parameters only through these auxiliary parameters and the risk aversion level R.

Here b_1 plays the role of a 'normalised discount factor', which adjusts the discount factor to allow for numéraire growth effects and for investment opportunities in the transaction-cost-free risky asset. The quantity b_4 is a simple function of the 'idiosyncratic volatility' of the illiquid asset. The parameter b_3 is the 'effective Sharpe ratio, per unit of idiosyncratic volatility' of the illiquid asset. The parameter b_2 is the hardest to interpret: essentially, it is a nonlinearity factor which arises from the multi-dimensional structure of the problem. Note that $b_2 = 1 + \frac{1}{1-\rho^2}(\frac{\beta}{\eta R} - \rho)^2 \geqslant 1$.

In the sequel, we work with the following assumption.

Standing Assumption 2.2 Throughout the paper, we assume $b_3 > 0$ and $b_2 > 1$.

The assumption $b_3 > 0$ is made only for convenience. However, the advantage of working with a positive effective Sharpe ratio of the illiquid asset ($b_3 > 0$) is that the no-transaction wedge is contained in the first two quadrants of the $(x, y\theta)$ plane. The assumption $b_3 > 0$ reduces the number of cases to be considered in our analysis, and facilitates the clarity of the exposition, but the methods and results developed in this

paper can be easily extended to the case of an illiquid asset with negative effective Sharpe ratio.[4]

The case $b_2 = 1$ is rather special and we exclude it from our analysis. One scenario in which we naturally find $b_2 = 1$ is if $\beta = 0 = \rho$. In this case, there is neither a hedging motive nor an investment motive for holding the liquid risky asset. Essentially, then, the investor can ignore the presence of the liquid risky asset, reducing the dimensionality of the problem. This problem is the subject of [14]. More generally, if $b_2 = 1$ (i.e., $\beta = R\eta\rho$) but $\beta \neq 0 \neq \rho$, then the presence of the liquid asset does impact the problem. Typically, the position in the liquid asset S is a combination of an investment position to take advantage of the expected excess returns in S and a hedging position to offset the risk of the position in the illiquid asset Y. If $\frac{\beta}{\eta R} = \rho$, then when $X = 0$, these terms exactly cancel. In particular, if the half-line $\{X = 0\}$ is inside the no-transaction region, then since consumption takes place from the cash account, if ever $X = 0$, then wealth can only go negative. Then the subspace $\{X \leq 0\}$ is absorbing, and no further purchases of the liquid asset are ever made, and the transaction cost on purchases becomes irrelevant. Mathematically, this is reflected in the fact that if $b_2 = 1$, then the solution n we define in the next section may pass through singular points. See Choi et al. [6] or Hobson et al. [14] for a discussion of some of the issues. A full analysis of the case $b_2 = 1$ requires a combination of the techniques in [6] and [5] or [14] and this paper, but for reasons of space, we do not pursue this here.

Remark 2.3 Let $\mathcal{P}_S$ be the Merton consumption/investment problem with the single risky asset S (and no transaction costs). Let $\mathcal{P}_Y^\xi$ be the Merton consumption/investment problem with the single risky asset Y and round-trip transaction cost ξ. Finally, let $\mathcal{P}_{S,Y}^\xi$ be the Merton consumption/investment problem for the pair of risky assets S, Y and round-trip transaction cost ξ as described in this section. In particular, let $\mathcal{P}_{S,Y}^0$ be the Merton consumption/investment problem with two risky assets and no transaction costs. The solutions to $\mathcal{P}_S$ and $\mathcal{P}_{S,Y}^0$ are well known; the solution to $\mathcal{P}_Y^\xi$ is given in Choi et al. [6] and Hobson et al. [14].

The necessary and sufficient condition for $\mathcal{P}_S$ to be well-posed is given by $b_1 > 0$. (If $R < 1$ and $b_1 \leq 0$, then there is an admissible strategy which yields infinite discounted expected utility of consumption. This strategy typically defers consumption from the present to the future, in order to build up large wealth reserves. If $R > 1$ and $b_1 \leq 0$, then every admissible strategy yields a discounted expected utility of consumption of minus infinity.) The necessary and sufficient condition for $\mathcal{P}_{S,Y}^0$ to be well-posed is $b_1 > \frac{1-R}{4R}b_3^2$.

If $R < 1$ and the problem $\mathcal{P}_S$ is ill-posed, then necessarily $\mathcal{P}_{S,Y}^\xi$ is ill-posed. An investor may simply liquidate his initial position in Y and then follow any admissible strategy for the single-asset problem involving S alone which yields infinite expected discounted utility. (If necessary, we may consider a sequence of admissible strategies involving investing in S alone.) In this case, $b_1 > 0$ is a necessary condition for $\mathcal{P}_{S,Y}^\xi$ to be well-posed.

---

[4]If $b_3 = 0$, the agent chooses never to invest in the illiquid asset. In this case, the agent closes any initial position in Y at time zero, and thereafter the problem reduces to a standard Merton problem with the single risky asset S and no transaction costs.
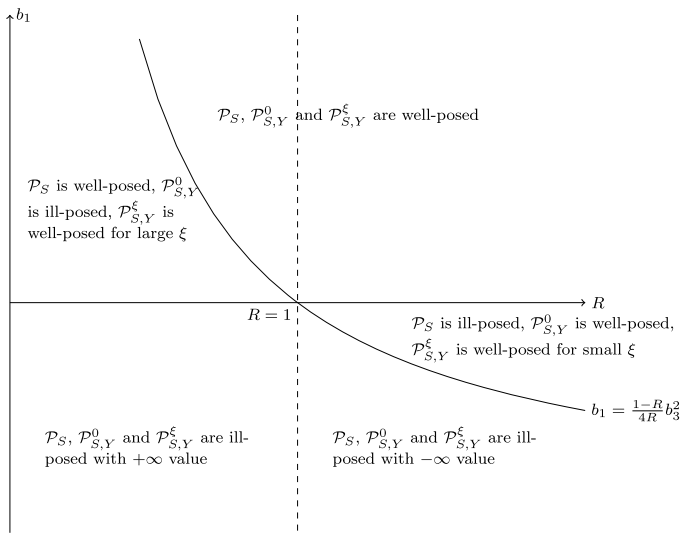
Fig. 1 All possible behaviours of the standard single-asset Merton problem ($\mathcal{P}_S$), the two-asset Merton problem without transaction costs ($\mathcal{P}_{S,Y}^0$) and the two-asset Merton problem with transaction costs ($\mathcal{P}_{S,Y}^\xi$), under different combinations of R and b_1 while keeping b_3 fixed

If $R > 1$ and the problem $\mathcal{P}_{S,Y}^0$ is ill-posed, then necessarily $\mathcal{P}_{S,Y}^\xi$ is ill-posed. An admissible strategy for the agent in the market with transaction costs is admissible for the problem $\mathcal{P}_{S,Y}^0$, but each such strategy yields value minus infinity. In this case, $b_1 > \frac{1-R}{4R}b_3^2$ is a necessary condition for $\mathcal{P}_{S,Y}^\xi$ to be well-posed.

Putting the two cases together, we find that a necessary condition for well-posedness is $b_1 > \min\{0, \frac{1-R}{4R}b_3^2\}$. If $b_1 > \max\{0, \frac{1-R}{4R}b_3^2\}$, we expect the value of $\mathcal{P}_{S,Y}^\xi$ to lie between the values of $\mathcal{P}_S$ and $\mathcal{P}_{S,Y}^0$, and since both these problems are well-posed, $\mathcal{P}_{S,Y}^\xi$ will be well-posed. More interestingly, if $\min\{0, \frac{1-R}{4R}b_3^2\} < b_1 \leq \max\{0, \frac{1-R}{4R}b_3^2\}$, then exactly one of $\mathcal{P}_S$ and $\mathcal{P}_{S,Y}^0$ is well-posed, and we expect the well-posedness of $\mathcal{P}_{S,Y}^\xi$ to depend on the level of transaction costs. In particular, if $R < 1$ and $0 < b_1 < \frac{1-R}{4R}b_3^2$, we expect $\mathcal{P}_{S,Y}^\xi$ to be well-posed only if transaction costs are sufficiently large; conversely, if $R > 1$ and $\frac{1-R}{4R}b_3^2 < b_1 < 0$, we expect $\mathcal{P}_{S,Y}^\xi$ to be well-posed only if transaction costs are sufficiently small. See Fig. 1 for a graphical illustration. This intuition is confirmed in Theorem 4.1.

Remark 2.4 Choi [5] assumes (in our notation) $\delta > r(1 - R)$ and $b_1 > \frac{(1-R)}{4R}b_3^2$. When $R < 1$, the latter assumption is precisely the condition under which $\mathcal{P}_{S,Y}^0$ is well-posed, and hence $\mathcal{P}_{S,Y}^\xi$ is well-posed for all levels of transaction costs. When $R > 1$, the assumption $\delta > r(1 - R)$ is a sufficient condition for the problem $\mathcal{P}_Y^\xi$ to be well-posed for every ξ, and hence $\mathcal{P}_{S,Y}^\xi$ is also well-posed for every ξ. Our work extends Choi [5] in the sense that we also consider the case where the well-posedness or otherwise of $\mathcal{P}_{S,Y}^\xi$ depends on the level of transaction costs.

Remark 2.5 In this paper, we assume there are just two risky assets, and that transaction costs are payable on just one of them. More generally, we may have several liquid assets on which no transaction costs are payable, provided the model includes no more than one risky asset which is subject to transaction costs. Although the equation for the dynamics of the wealth in liquid assets is then more complicated, the problem reduces to a univariate problem in which the state variable is the ratio of wealth in the asset on which transaction costs are payable to liquid wealth; moreover, after scaling for paper wealth, the HJB equation for the value function, see (3.3) below, is identical to the case of two risky assets which we study. See Evans et al. [11, Sect. 2.1] and Bichuch and Guasoni [2, Sect. 3.4] for a discussion of the multi-asset case and how it may simplify.

3 Heuristic derivation of a free boundary value problem

Inspired by the analysis in the classical case involving a single risky asset only, we postulate that the value function has the form

$$V(x, y, \theta) = \Upsilon \frac{(x + y\theta)^{1-R}}{1 - R} G\left(\frac{y\theta}{x + y\theta}\right) \tag{3.1}$$

for some strictly positive function G to be determined and $\Upsilon = b_4^R R^R$, a convenient scaling constant which helps to simplify the HJB equation.

Write $p := \frac{y\theta}{x+y\theta}$. As in the single-asset case (see Hobson et al. [14]), outside the no-transaction region, it is straightforward to deduce the form of G to be

$$G(p) = \begin{cases} A_*(1 + \lambda p)^{1-R}, & -\frac{1}{\lambda} \le p < p_*, \\ A^*(1 - \gamma p)^{1-R}, & p^* < p \le \frac{1}{\gamma}, \end{cases} \tag{3.2}$$

for some positive constants A_* and A^* to be determined.

Now we consider the no-transaction region. We expect the process

$$M_t := \int_0^t e^{-\delta s} \frac{C_s^{1-R}}{1 - R} ds + e^{-\delta t} V(X_t, Y_t, \Theta_t)$$

to be a supermartingale in general, and a martingale under the optimal strategy. Suppose V is $C^{2,2,1}$ and strictly increasing and concave in x. Using Itô's lemma and then maximising the drift term of M with respect to C and Π, we can formally obtain the HJB equation over the no-transaction region as

$$\frac{R}{1 - R} V_x^{1-1/R} + rx V_x + \alpha y V_y + \frac{\eta^2}{2} y^2 V_{yy} - \frac{(\beta V_x + \eta \rho y V_{xy})^2}{2V_{xx}} - \delta V = 0. \tag{3.3}$$

Our initial objective is to simplify (3.3). First, (3.1) can be used to reduce (3.3) to a second order, nonlinear equation for $G = G(p)$. Then, away from $p = 1$, we set[5]

---

[5]The assumption $b_3 > 0$ means that the agent would like to hold positive quantities of the illiquid asset, and that the no-transaction wedge is contained in the half-space $\{p > 0\}$. To allow $b_3 < 0$, it is necessary to consider $p < 0$. This case can be incorporated into the analysis by incorporating an extra factor of sgn(p) into the definition of h, so that $h(p) = \text{sgn}(p(1 - p))|1 - p|^{R-1}G(p)$. This then leads to extra cases, but no new mathematics, and the problem can still be reduced to solving $n' = O(q, n)$ where O is given by (3.6), but now for $q < 0$.

$h(p) = \text{sgn}(1-p)|1-p|^{R-1}G(p)$, $w(h) = p(1-p)\frac{dh}{dp}$, $W(h) = \frac{w(h)}{(1-R)h}$, $N = W^{-1}$ (the inverse of W) and $n(q) = |N(q)|^{-1/R}|1-q|^{1-1/R}$. Then[6] the HJB equation (3.3) can be reduced to

$$0 = \frac{n(q)}{b_4} - \delta + r(1-R)(1-q) + \alpha(1-R)q + \frac{\eta^2}{2}(1-R)\Big(qw'(N(q)) - q\Big)$$

$$- \frac{1-R}{2}\frac{(\beta(1-q) - \eta\rho(qw'(N(q)) - (1-R)q))^2}{qw'(N(q)) + (2R-1)q - R}. \tag{3.4}$$

Details of the algebra behind the transformation can be found in [14].

After multiplying through by the denominator of the last term, this can be viewed as a quadratic equation in $qw'(N(q))$. We want the root corresponding to $V_{xx} < 0$. This is equivalent to

$$\frac{1}{1-R}p^2 G''(p) + \frac{2R}{1-R}pG'(p) - RG(p) < 0, \tag{3.5}$$

which can be restated as $qw'(N(q)) + (2R-1)q - R < 0$.

From the relationships between w, W, N and n, we have

$$\frac{n'(q)}{n(q)} = \frac{1-R}{R(1-q)} - \frac{1}{R}\frac{N'(q)}{N(q)} = \frac{1-R}{R(1-q)} - \frac{1-R}{R}\frac{q}{qw'(N(q)) - (1-R)q^2}.$$

After some algebra, we arrive at the ODE $n'(q) = O(q, n(q))$, where

$$O(q,n) = \frac{(1-R)n}{R(1-q)} - \frac{2(1-R)^2 qn/R}{K(q, \varphi(q,n), E(q))} \tag{3.6}$$

with $E(q)^2 := 4R^2(1-R)^2(b_2-1)(1-q)^2$ and

$$\varphi(q,n) := n - b_1 + (1-R)(b_3 - 2R)q + (2-b_2)R(1-R),$$

$$K(q,\phi,E) := 2(1-R)(1-q)\big((1-R)q + R\big) - \phi - \text{sgn}(1-R)\sqrt{\phi^2 + E^2}. \tag{3.7}$$

Note that the sign in front of the square root in (3.7) comes from the choice of root for $qw'(N(q))$ in (3.4) corresponding to concavity of V.

Define the quadratic function

$$m(q) := R(1-R)q^2 - b_3(1-R)q + b_1$$

and the algebraic function

$$\ell(q) := m(q) + (1-R)q(1-q) + (b_2-1)R(1-R)\frac{q}{(1-R)q + R}.$$

---

[6]The key transformation is to set $w(h) = p(1-p)\frac{dh}{dp}$. This has the effect of changing the independent variable, and more importantly reducing the equation to a first order equation.

Note that m has a turning point (a minimum if $R < 1$ and a maximum if $R > 1$) at $q_M := \frac{b_3}{2R}$, and set $m_M := m(q_M) = b_1 - \frac{b_3^2(1-R)}{4R}$. The lemma below gives the key properties of O.

Lemma 3.1 1) *$O(q,n)$ can be extended to $q = 1$ on $\{(1-R)n < (1-R)\ell(1)\}$ by continuity.*

2) *$O(q,n) = 0$ if and only if $n = m(q)$.*

3) *For given R and q, the sign of $O(q,n)$ depends only on the signs of $n - m(q)$ and $\ell(q) - n$.*

Now we apply the same transformations on the purchase and sale region. For $-\frac{1}{\lambda} \leqslant p < p_*$, we have $G(p) = A_* (1 + \lambda p)^{1-R}$ as given by (3.2). Then

$$w(h) = p(1-p)\frac{dh}{dp} = p(1-p)(1-R)h\left(\frac{\lambda}{1+\lambda p} + \frac{1}{1-p}\right)$$

$$= (1-R)h\left(\frac{p(1+\lambda)}{1+\lambda p}\right)$$

and $|1 - W(h)| = \frac{|1-p|}{1+\lambda p} = \left(\frac{A_*}{|h|}\right)^{1/(1-R)}$. It follows that $n(q) = (A_*)^{-1/R}$. This expression holds for $-\frac{1}{\lambda} \leqslant p < p_*$, for which $q = W(h) = \frac{(1+\lambda)p}{1+\lambda p}$. The equivalent range in q is thus given by $q < q_* := \frac{(1+\lambda)p_*}{1+\lambda p_*}$. Similarly, on the sale region, we have $n(q) = (A^*)^{-1/R}$ for $q > q^* := \frac{(1-\gamma)p^*}{1-\gamma p^*}$.

The $C^{2,2,1}$ smoothness of the original value function V translates into C^1 smoothness of the transformed value function n. Hence we are looking for a positive, continuously differentiable function n and boundary points (q_*, q^*) solving $n' = O(q,n)$ on $\{q \in (q_*, q^*)\}$ with $n(q) = (A_*)^{-1/R}$ for $q \leq q_*$ and $n(q) = (A^*)^{-1/R}$ for $q \geq q^*$. First order smoothness of n at the boundary points forces that we must have $n'(q_*) = n'(q^*) = 0$. But by Lemma 3.1, $n'(q) = O(q, n(q)) = 0$ if and only if $n(q) = m(q)$. Hence the free boundary points must be given by the q-coordinates where n intersects the quadratic m. The free boundary value problem now becomes solving $n'(q) = O(q, n(q))$ on $\{q \in (q_*, q^*)\}$ subject to $n(q_*) = m(q_*)$ as well as $n(q^*) = m(q^*)$.

Recall the definition of the round-trip transaction cost $\xi = \frac{\lambda+\gamma}{1-\gamma} > 0$. Suppose for now that $1 \notin [p_*, p^*]$ and in turn $1 \notin [q_*, q^*]$. Exploiting the relationships $q_* = \frac{(1+\lambda)p_*}{1+\lambda p_*}$ and $q^* = \frac{(1-\gamma)p^*}{1-\gamma p^*}$, we have

$$\ln(1+\xi) = \ln(1+\lambda) - \ln(1-\gamma) = \int_{p_*}^{p^*} \frac{dp}{p(1-p)} - \int_{q_*}^{q^*} \frac{dq}{q(1-q)}.$$

Then, applying the definitions of w, N and O and using

$$\int_{p_*}^{p^*} \frac{dp}{p(1-p)} = \int_{h_*}^{h^*} \frac{dh}{w(h)} = \int_{q_*}^{q^*} \frac{N'(q)\,dq}{(1-R)q\,N(q)}$$

and

$$\frac{O(q, n(q))}{n(q)} = \frac{n'(q)}{n(q)} = \frac{1 - R}{R(1 - q)} - \frac{1}{R} \frac{N'(q)}{N(q)},$$

we have

$$\ln(1 + \xi) = \int_{q_*}^{q^*} \left(-\frac{R}{q(1 - R)} \frac{O(q, n(q))}{n(q)} \right) dq. \tag{3.8}$$

Hence the required solution from the free boundary value problem is the one such that (3.8) holds.

If $1 \in [p_*, p^*]$ or equivalently $1 \in [q_*, q^*]$, the two integrals $\int_{p_*}^{p^*} \frac{dp}{p(1-p)}$ and $\int_{q_*}^{q^*} \frac{dq}{q(1-q)}$ are not well defined. But it can be shown that (3.8) still holds by using a limiting argument.

To summarise, we are interested in solving the following *free boundary value problem*.

Problem 3.2 Find a positive function $n(\cdot)$ and a pair of boundary points (q_*, q^*) solving

$$n'(q) = O(q, n(q)), \qquad q \in [q_*, q^*], \tag{3.9}$$

$$n(q_*) = m(q_*), \qquad n(q^*) = m(q^*), \tag{3.10}$$

subject to (3.8).

In Sect. 5, we distinguish several different cases and discuss how to construct the solution $(n(\cdot), q_*, q^*)$ in each of these cases. The switch of coordinate from $p = \frac{y\theta}{x+y\theta}$ to q is, at one level, purely a mathematical simplification. But the quantities q_*, q^* indeed have an important interpretation as the critical proportion of wealth invested in the illiquid asset when the position is evaluated using the ask or bid price. We shall also see in Sect. 5 that the values of $m_M = m(q_M)$ and $\max_{0 \le q \le 1} \frac{\ell(q)}{R-1}$ are crucial in determining when the problem is ill-posed.

4 Main results

Given a solution $(n(\cdot), q_*, q^*)$ to Problem 3.2, we can reverse the transformations in Sect. 3 and construct a candidate value function. The key technicality here is to demonstrate that the function we construct is smooth up to the second order and satisfies a HJB variational inequality. These issues are covered by Appendices A and B.

The first pair of main results of this paper are summarised in the following two theorems. For a given set of risk aversion parameter R, discount factor δ and market parameters r, μ, σ, α, η, ρ, we say the problem is *unconditionally well-posed* if the value function is finite on the interior of the solvency region for all values of the transaction costs $\lambda \ge 0$ and $\gamma \in [0, 1)$ with $\lambda + \gamma > 0$. We say the problem is *ill-posed*

if the value function is infinite for all λ and γ. We say the problem is *conditionally well-posed* if the problem is well-posed for some values of the round-trip transaction cost, but ill-posed for other values.

Let $L(q) = \frac{b_1 - \ell(q)}{1-R} = (b_3 - 1)q + (1-R)q^2 - \frac{(b_2-1)Rq}{(1-R)q+R}$ and define the constant $L^* = \max_{0 \le q \le 1} L(q)$. Note that L and L^* depend on R, b_2 and b_3, but not on b_1. For $R < 1$, L is convex on $[0, 1]$ and $L^* = \max\{L(0), L(1)\} = (b_3 - b_2 R)^+$. If $R > 1$, then L is concave and there is no correspondingly simple expression for L^*, although we have the simple bound $L^* \ge (b_3 - b_2 R)^+$.

Theorems 4.1 and 4.3 are proved in Appendix C.

Theorem 4.1 *The investment/consumption problem* (*Problem* 2.1) *is well-posed if and only if there is a positive solution to Problem* 3.2. *In particular, it is*
1) *unconditionally well-posed if*
    (a) $R < 1$ *and* $b_1 \ge \frac{1-R}{4R} b_3^2$;
    (b) $R > 1$ *and* $b_1 \ge -(R-1)L^*$;
2) *ill-posed if*
    (a) $R < 1$ *and* $b_1 \le (1-R)(b_3 - b_2 R)^+$;
    (b) $R > 1$ *and* $b_1 \le \frac{1-R}{4R} b_3^2$;
3) *conditionally well-posed if*
    (a) $R < 1$ *and* $(1-R)(b_3 - b_2 R)^+ < b_1 < \frac{1-R}{4R} b_3^2$; *in this case, the problem is well-posed if and only if* $\xi > \bar{\xi}$, *where* $\bar{\xi}$ *is defined in* (5.2) *below*;
    (b) $R > 1$ *and* $\frac{1-R}{4R} b_3^2 < b_1 < -(R-1)L^*$; *in this case, the problem is well-posed if and only if* $\xi < \bar{\xi}$.

Note that $\frac{b_3^2}{4R} \ge b_3 - R > b_3 - b_2 R$ so that in the case $R < 1$, we have the inequality $\frac{1-R}{4R} b_3^2 > (1-R)(b_3 - b_2 R)$.

The condition $b_1 > (1-R)(b_3 - b_2 R)$ simplifies to $\delta > (1-R)(\alpha - \frac{\eta^2 R}{2})$. Given the results of [14] on the single-asset case, we have the following corollary.

Corollary 4.2 *Suppose* $R < 1$. *The problem with risky liquid asset and illiquid risky asset is ill-posed* (*for all values of transaction costs*) *if and only if the problem with the liquid asset alone is ill-posed or the problem with risky liquid asset omitted is ill-posed* (*for all values of transaction costs*).

When $R < 1$, it is clear that if $\mathcal{P}_S$ is ill-posed or if $\mathcal{P}_Y^\xi$ is ill-posed, then $\mathcal{P}_{S,Y}^\xi$ is ill-posed; so the main content of Corollary 4.2 is the 'only if' statement.

The next theorem links the solution of the free boundary value problem and that of the optimal investment/consumption problem.

Theorem 4.3 *Suppose the parameters are such that Problem* 2.1 *is well-posed and* $(n(\cdot), q_*, q^*)$ *is the solution to the free boundary value problem* (*Problem* 3.2). *Set*

$$V^C(x, y, \theta) = \left(\frac{b_1}{Rb_4}\right)^{-R} \frac{(x + y\theta)^{1-R}}{1-R} G^C\left(\frac{y\theta}{x + y\theta}\right),$$

where G^C is a C^2 function constructed from n as described in Proposition A.1 of Appendix A. Then $V^C = V$ where V is the value function of the investment/consumption problem defined in (2.1). The purchase and sale boundaries of the illiquid asset are given by

$$p_* = \frac{q_*}{1 + \lambda(1 - q_*)}, \qquad p^* = \frac{q^*}{1 - \gamma(1 - q^*)}. \tag{4.1}$$

Purchase of the illiquid asset occurs whenever $\frac{y\theta}{x+y\theta} = p < p_*$. Using (4.1), this condition can be rewritten as $\frac{y\theta(1+\lambda)}{x+y\theta(1+\lambda)} < q_*$. Hence q_* can also be viewed as a critical threshold at which purchase occurs, but now the illiquid asset is valued at the ask price $y(1+\lambda)$ instead of the pre-transaction cost price y. A similar interpretation holds for q^*.

5 Solutions to the free boundary value problem

In this section, we discuss the key features of the solutions to Problem 3.2.

Recall that (q_M, m_M) is the extreme point of the quadratic m (a minimum when $R < 1$ and a maximum when $R > 1$) with $q_M = \frac{b_3}{2R} > 0$. The key analytical properties of the problem only depend on the signs of the four parameters b_1, $1 - R$, m_M, $\max_{0 \le q \le 1}(R - 1)\ell(q)$. We classify six different cases. In the analysis of the cases, we make extensive use of the properties of O given in Lemma 3.1 and Lemma D.1 of Appendix D.

We parametrise the family of solutions to (3.9) by the left boundary point. Fix u such that $m(u) \ge 0$ and denote by $(n_u(q))_{q \ge u}$ the solution to the initial value problem

$$n'(q) = O(q, n(q)), \qquad n(u) = m(u).$$

Let $\zeta(u) = \inf\{q \ge u : (1 - R)n_u(q) < (1 - R)m(q)\}$ denote where n_u first crosses m to the right of u. Define

$$\Sigma(u) = \exp\left(\int_u^{\zeta(u)} \left(-\frac{R}{q(1-R)} \frac{O(q, n_u(q))}{n_u(q)}\right) dq\right) - 1. \tag{5.1}$$

Let $F(q, n) = \frac{O(q,n)}{n}$ and set $F(q, 0) = \lim_{n \downarrow 0} F(q, n)$. Let $p_- \le p_+$ be the roots of $m(q) = 0$. Set

$$\overline{\xi} := \exp\left(-\int_{p_-}^{p_+} \frac{R}{q(1-R)} F(q, 0) dq\right) - 1. \tag{5.2}$$

Lemma 5.1 1) *Suppose $R < 1$.*

(a) *Suppose $b_1 \ge \frac{1-R}{4R} b_3^2$. Then Σ is a strictly decreasing, continuous mapping $\Sigma : (0, q_M] \to [0, \infty)$ with $\Sigma(0+) = +\infty$ and $\Sigma(q_M) = 0$.*

(b) *Suppose $(1 - R)(b_3 - b_2 R)^+ < b_1 \le \frac{1-R}{4R} b_3^2$. Then Σ is a strictly decreasing, continuous mapping $\Sigma : (0, p_-] \to [\overline{\xi}, \infty)$ with $\Sigma(0+) = +\infty$ and $\Sigma(p_-) = \overline{\xi}$. Moreover, $\lim_{u \uparrow p_-} n_u(\cdot) = 0$ and $\lim_{u \uparrow p_-} \zeta(u) = p_+$.*

2) *Suppose $R > 1$.*

(a) *Suppose $b_1 \geq -L^*(R-1)$. Then Σ is a strictly decreasing, continuous mapping $\Sigma : (p_-, q_M] \to [0, \infty)$ with $\Sigma(p_-+) = +\infty$ and $\Sigma(q_M) = 0$.*

(b) *Suppose $\frac{1-R}{4R} b_3^2 < b_1 < -L^*(R-1)$. Then Σ is a strictly decreasing, continuous mapping $\Sigma : (p_-, q_M] \to [0, \overline{\xi})$ with $\Sigma(p_-) = \overline{\xi}$ and $\Sigma(q_M) = 0$. Moreover, $\lim_{u \downarrow p_-} n_u(\cdot) = 0$ and $\lim_{u \downarrow p_-} \zeta(u) = p_+$.*

Lemma 5.1 is proved in Appendix D.

5.1 The cases

5.1.1 Case 1: $R < 1$ and $m_M \geq 0$. Equivalently, this may be stated as $R < 1$ and $b_1 \geq \frac{1-R}{4R} b_3^2$

For any initial value $u \in (0, q_M)$,

$$m'(u) < 0 = O\big(u, m(u)\big) = O\big(u, n_u(u)\big) = n_u'(u).$$

Thus $n_u(q)$ must initially be larger than $m(q)$ for q being close to u. It can be checked that $O(q, n)$ is negative on the set

$$\{(q, n) : 0 < q \leqslant 1, m(q) < n < \ell(q)\} \cup \{(q, n) : q > 1, n > m(q)\}$$

(see part 4 of Lemma D.1 in Appendix D). Also, $n_u(q)$ cannot cross $\ell(q)$ from below on $\{0 < q \leqslant 1\}$ since $\lim_{n \uparrow \ell(q)} O(q, n) = -\infty$; see for example (D.4). By considering the sign of $O(q, n)$, we conclude that n_u must be decreasing until it crosses m. This guarantees the finiteness of $\zeta(u)$, and the triple $(n_u(\cdot), u, \zeta(u))$ represents one possible solution to problem (3.10). Notice that the family of solutions $(n_u(\cdot))_{0 < u < q_M}$ cannot cross, and thus $n_u(q)$ is decreasing in u. The solutions corresponding to initial values $u = 0$ and $u = q_M$ can be understood as the appropriate limits of a sequence of solutions.

Although $O(q, n)$ has singularities at $q = 1$ and $n = \ell(q)$, a well-defined limit $O(q, n)$ exists on $\{(q, n) : q = 1, n < \ell(1)\}$ and $\{(q, n) : q > 1, n = \ell(q)\}$ (see part 3 of Lemma D.1 in Appendix D). Hence there exists a continuous modification of $O(q, n)$, and a solution n_u can actually pass through these singularity curves. See Fig. 2(a) for some examples.

From the analysis leading to (3.8), the correct choice of u should satisfy $\xi = \Sigma(u)$. From Lemma 5.1, for every given level of round-trip transaction cost ξ, there exists a unique choice of the left boundary point given by $u_* = \Sigma^{-1}(\xi)$, and then the desired solution to the free boundary value problem is given by $(n_{u_*}(\cdot), u_*, \zeta(u_*))$. Figure 2(b) gives the plots of $\Sigma^{-1}(\xi)$ and $\zeta(\Sigma^{-1}(\xi))$ representing the boundaries (q_*, q^*) under different levels of transaction costs.

5.1.2 Case 2: $R < 1$ and either $b_1 \leq 0$ or $b_1 > 0$, $m_M < 0$, $\ell(1) \leq 0$. Equivalently, this may be stated as $R < 1$ and $b_1 \leq (1 - R)(b_3 - b_2 R)^+$

If $b_1 \leq 0$, then m is negative on $(0, q_M]$ and there can be no nonnegative solutions to $n' = O(q, n)$ with $n(u) = m(u)$ for $u \in (0, q_M]$.
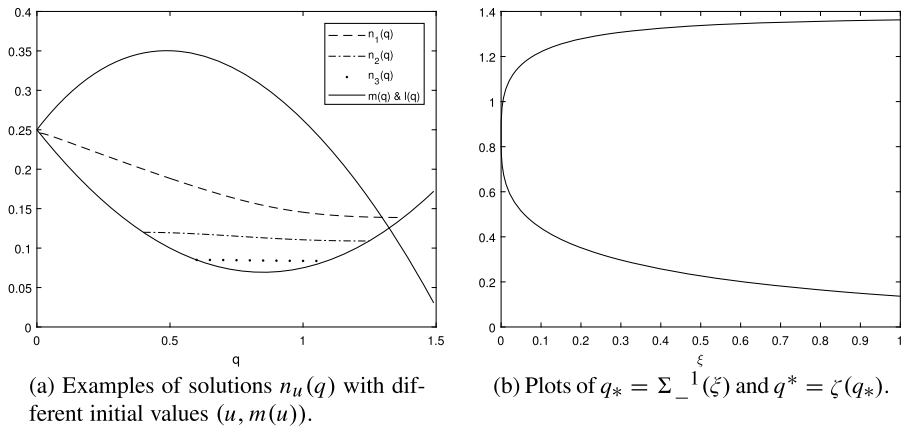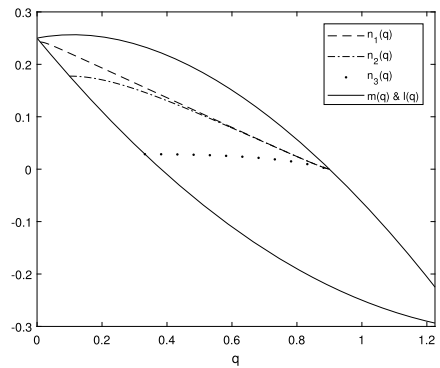
(a) Examples of solutions $n_u(q)$ with different initial values $(u, m(u))$.

(b) Plots of $q_* = \Sigma^{-1}(\xi)$ and $q^* = \zeta(q_*)$.

Fig. 2 Case 1. Parameters are $R = 0.5$, $b_1 = 0.25$, $b_2 = 1.75$, $b_3 = 0.85$

Fig. 3 Case 2. Parameters are $R = 0.5$, $b_1 = 0.25$, $b_2 = 1.75$, $b_3 = 1.5$. Then $q_M = \frac{b_3}{2R} = 1.5$

Let p^ℓ be the root of $\ell(q) = 0$ on $\{q \in (0, 1)\}$. Since the solution to the ODE $n'(q) = O(q, n(q))$ must be bounded below by zero and above by $\ell(q)$ on the range $\{q \in (0, p^\ell)\}$ for any initial value $(u, m(u))$ for which $m(u) > 0$, the corresponding solution $n_u(\cdot)$ must hit $(p^\ell, 0)$; see Fig. 3. Hence there does not exist any positive solution which crosses m again to the right of u, and there is no solution to the free boundary value problem. Indeed, under the combination of parameters in this case, it follows from part (b) of Theorem 4.1 that the underlying problem is ill-posed for all levels of transaction costs and thus the value function cannot be defined.

5.1.3 Case 3: $R < 1$ and $b_1 > 0$, $m_M < 0$, $\ell(1) > 0$. Equivalently, this may be stated as $R < 1$ and $(1 - R)(b_3 - b_2 R)^+ < \frac{1-R}{4R}b_3^2$

Let $p_\pm$ with $0 < p_- < q_M < p_+$ be the two roots of $m(q) = 0$. The parametrisation of the solution is the same as in Case 1 except that the left boundary point should now be restricted to $\{u \in (0, p_-)\}$ to ensure a positive initial value. The function Σ defined in (5.1) is still a strictly decreasing map with $\Sigma(0+) = +\infty$ except that its domain is now restricted to $(0, p_-]$.

(a) Examples of solutions $n_u(q)$ with different initial values $(u, m(u))$.

(b) Plots of $q_* = \Sigma^{-1}(\xi)$ and $q^* = \zeta(q_*)$. q_* and q^* are not defined for $\xi < \bar{\xi}$.
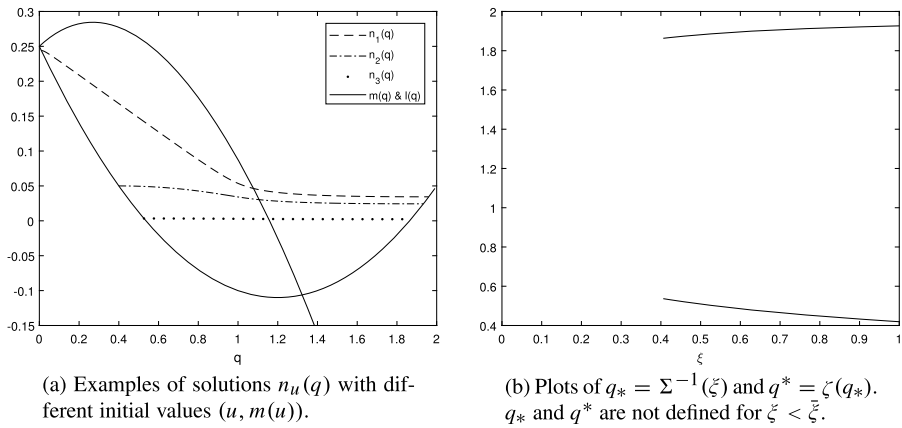
Fig. 4 Case 3. Parameters are $R = 0.5$, $b_1 = 0.25$, $b_2 = 1.75$, $b_3 = 1.2$

Unlike Case 1, we now only consider $\Sigma^{-1}(\xi)$ on the range $\{\xi \in (\bar{\xi}, \infty)\}$. For such a given high level of round-trip transaction cost, the required left boundary point is given by $u_* = \Sigma^{-1}(\xi)$ and $u^* = \zeta(u_*)$; see Fig. 4. In this case, the problem is conditionally well-posed, i.e., it is well-posed only for a sufficiently high level of transaction cost.

5.1.4 Case 4: $R > 1$ and either $b_1 \geq 0$ or both $b_1 < 0$ and $\max_{0 \leq q \leq 1} \ell(q) \geq 1$. Equivalently, this may be stated as $R > 1$ and $b_1 \geq -L^*(R - 1)$

Suppose first $b_1 \geq 0$. In this case, the quadratic m is positive on $(0, q_M]$, has a positive maximum at (q_M, m_M) and $m(q) > \ell(q)$ on $\{q \in (0, 1)\}$. By checking the sign of $O(q, n)$ in case of $R > 1$, one can verify that the solution n_u of the initial value problem is always increasing for any choice of left boundary point $u \in (0, q_M)$. In this case, the family of solutions is increasing in u. The solution $n_u(q)$ crosses $m(q)$ from below at $\zeta(u) = \inf\{q \geqslant u : n_u(q) > m(q)\}$. The correct choice of u is again the one solving $\xi = \Sigma(u)$, where Σ is defined in (5.1). As in Case 1, the function Σ is onto from $(0, q_M]$ to $[0, \infty)$, and hence $u_* = \Sigma^{-1}(\xi)$ always exists uniquely for any ξ; see Fig. 5.

Now suppose that $-L^*(R - 1) \leq b_1 < 0$, from which it follows that ℓ has a root $p^\ell \in (0, 1]$. As before, we can define a nonnegative, increasing solution n_u for any $u \in (p_-, q_M)$ which crosses $m(q)$ from below at $\zeta(u)$. This family of solutions is increasing in u and each member is bounded below by $\ell(q)$. Let n_{p_-} be given by $n_{p_-}(q) = \lim_{u \downarrow p_-} n_u(q)$. Then $n_{p_-}(q) = 0$ for $p_- < q \leq p^\ell$, but $n_{p_-}(q) > \ell(q)$ for $p^\ell < q \leq 1$. By Lemma 5.1, Σ is onto from $(p_-, q_M]$ to $[0, \infty)$ and hence $u_* = \Sigma^{-1}(\xi)$ always exists for any ξ.

5.1.5 Case 5: $R > 1$ and $m_M \leq 0$. Equivalently, this may be stated as $b_1 \leq -\frac{R-1}{4R}b_3^2$

These parameter values are such that m is nonpositive on $(0, q_M]$ and there can be no positive solution to $n' = O(q, n)$ with initial condition $n(u) = m(u)$.
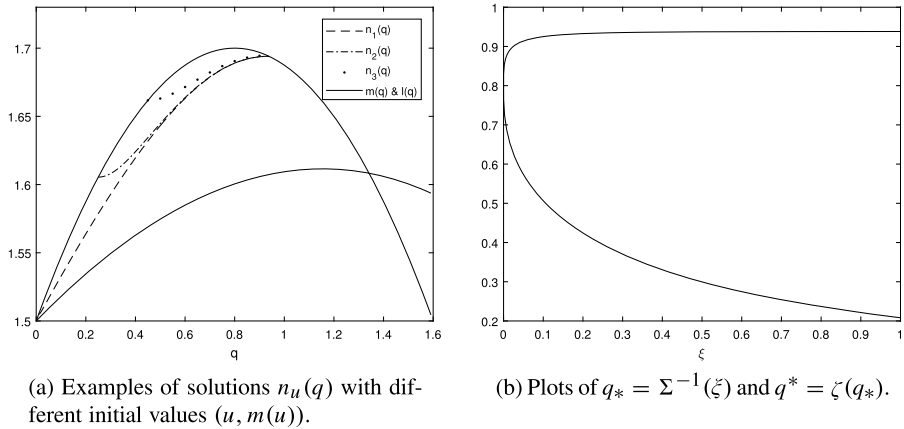
(a) Examples of solutions $n_u(q)$ with different initial values $(u, m(u))$.

(b) Plots of $q_* = \Sigma^{-1}(\xi)$ and $q^* = \zeta(q_*)$.

Fig. 5 Case 4. Parameters are $R = 1.25$, $b_1 = 1.5$, $b_2 = 1.25$, $b_3 = 2$

(a) Examples of solutions $n_u(q)$ with different initial values $(u, m(u))$.

(b) Plots of $q_* = \Sigma^{-1}(\xi)$ and $q^* = \zeta(q_*)$. q_* and q^* are not defined for $\xi > \bar{\bar{\xi}}$.

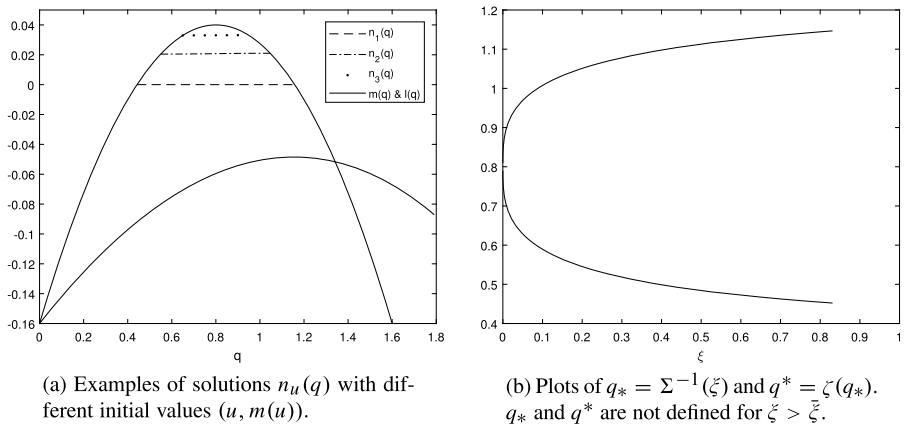**Fig. 6** Case 6. Parameters are $R = 1.25$, $b_1 = -0.16$, $b_2 = 1.25$, $b_3 = 2$

5.1.6 Case 6: $R > 1$, $b_1 < 0$, $m_M > 0$ and $\max_{0 \leq q \leq 1} \ell(q) < 0$. *Equivalently, this may be stated as $R > 1$ and $-\frac{R-1}{4R} b_3^2 < b_1 < -L^*(R-1)$*

Since $b_1 < 0$, the solutions n_u are only relevant for $u \in (p_-, q_M)$. Over this interval, solutions n_u are increasing in u. Further,

$$n_{p_-}(q) := \lim_{u \downarrow p_-} n_u(q) = 0 > \ell(q)$$

on $(0, 1)$. For $q > 1$, the differential equation is such that n may cross ℓ and hence we find $n_{p_-}(q) = 0$ for $q \in [p_-, p_+]$. It follows that for $\xi < \bar{\bar{\xi}}$, there is a positive solution to Problem 3.2, but for $\xi \geq \bar{\bar{\xi}}$, there is no positive solution; see Fig. 6.

Based on the figures in the various cases, we can make a series of simple observations about the behaviour of q_* and q^*. Some of these are proved in Sect. 6 on

the comparative statics of the problem. First, the lower and upper boundaries of the no-transaction region, expressed in terms of q_* and q^*, are monotonically decreasing and monotonically increasing, respectively, in ξ. In particular, the no-transaction region gets wider as transaction costs increase. Second, the no-transaction region may be contained in the first quadrant $(0 < q_* < q^* < 1)$, or the upper half-plane $(0 < q_* < 1 < q^*)$, or the second quadrant $(1 < q_* < q^*)$. Third, when the problem is well-posed for small transaction costs, $\lim_{\xi \downarrow 0} q_* = q_M = \lim_{\xi \downarrow 0} q^*$. Moreover, when the problem is well-posed for large transaction costs, $\lim_{\xi \uparrow \infty} q_* = 0$ and $\lim_{\xi \uparrow \infty} q^* =: q^*_\infty < \infty$ so that there is a part of the solvency space close to the solvency limit $p = 1/\gamma$ which even in the regime of very large transaction costs is inside the region where a sale of Y at $t = 0$ is necessary. Fourth, q^* is generally less sensitive to changes in ξ than q_* so that the no-transaction wedge is not centred on the Merton line; we have $q^*(\xi) - q_M < q_M - q_*(\xi)$. However, in cases where the problem is conditionally well-posed, $\lim_{\xi \to \bar{\xi}} q^*(\xi) - q_M = q_M - \lim_{\xi \to \bar{\xi}} q^*(\xi)$, and in the limit, the no-transaction wedge is centred around the Merton line.

6 Comparative statics

In this section, we investigate how the no-transaction wedge $[p_*, p^*]$ and the value function V change with the market parameters and the level of transaction costs.

6.1 Monotonicity with respect to market parameters

Proposition 6.1 *Suppose $(n(\cdot), q_*, q^*)$ is the solution to the free boundary value problem* (*Problem* 3.2). *Then*:

1) *q_* and q^* are decreasing in b_1.*
2) *For $R < 1$, q_* and q^* are increasing in b_3.*

Proposition 6.1 is proved in Appendix E.

Recall that $p_* = \frac{q_*}{1 + \lambda(1 - q_*)}$ and $p^* = \frac{q^*}{1 - \gamma(1 - q^*)}$. Then Proposition 6.1 immediately gives

Theorem 6.2 1) *p_* and p^* are decreasing in b_1.*
2) *For $R < 1$, p_* and p^* are increasing in b_3.*

Theorem 6.2 describes the comparative statics in terms of the auxiliary parameters.[7] In general, it is difficult to make categorical statements about the comparative statics with respect to the original market parameters since many of these enter the definitions of more than one of the auxiliary parameters. However, we have the following results concerning the dependence of p_* and p^* on the discount rate and on the drift of the illiquid asset.

---

[7]Since the free boundary value problem does not depend on b_4, q_* and q^* are trivially independent of b_4. We have strong numerical evidence that q_* is decreasing in b_2 and q^* is increasing in b_2, but we have not been able to prove this result.

Corollary 6.3 p_* and p^* are decreasing in δ. If $R < 1$, then p_* and p^* are increasing in α.

The corollary confirms the intuition that as the return on the illiquid asset increases, it becomes more valuable and the agent elects to buy the illiquid asset sooner, and to sell it later. Moreover, as his discount parameter increases, he wants to consume wealth sooner, and since consumption takes place from the cash account, he elects to keep more of his wealth in liquid assets, and less in the illiquid asset.

Now we consider the equivalent cash value of the holdings in the illiquid asset. We compare the agent with holdings in the illiquid asset to an otherwise identical agent (same risk aversion and discount parameter, and trading in the financial market with bond and risky asset with price S) who has a zero initial endowment in the illiquid asset and is precluded from taking any positions in the illiquid asset.

Consider the market without the illiquid asset. For an agent operating in this market, a consumption/investment strategy is admissible for initial wealth $x > 0$ (we write $(C = (C_t)_{t \geq 0}, \Pi = (\Pi_t)_{t \geq 0}) \in \mathcal{A}_W(x)$) if C and Π are progressively measurable and the resulting wealth process $X = (X_t)_{t \geq 0}$ is nonnegative for all t. Here X solves $dX_t = r(X_t - \Pi_t)\,dt + \frac{\Pi_t}{S_t}dS_t - C_t\,dt$ subject to $X_0 = x$. Let $W = W(x)$ be the value function for a CRRA investor, i.e.,

$$W(x) = \sup_{(C,\Pi) \in \mathcal{A}_W(x)} \mathbb{E}\left[\int_0^\infty e^{-\delta t}\frac{C_t^{1-R}}{1-R}\,dt\right].$$

The problem of finding W is a classical Merton consumption/investment problem without transaction costs. For the problem to be well-posed, we require that $b_1 > 0$. **For the rest of this section, we assume that $b_1 > 0$.** We find $W(x) = (\frac{b_1}{b_4 R})^{-R}\frac{x^{1-R}}{1-R}$.

Define $\mathcal{C} = \mathcal{C}(y\theta; x)$ to be the certainty equivalent value of holding the illiquid asset, i.e., the cash amount which the agent with liquid wealth x and θ units of illiquid asset with current price y, trading in the market with transaction costs, would exchange for his holdings of the illiquid asset if after this exchange he is not allowed to trade in the illiquid asset. (We assume there are no transaction costs on this exchange, but they can be easily added if required.) Then $\mathcal{C} = \mathcal{C}(y\theta; x)$ solves $W(x + \mathcal{C}) = V(x, y, \theta)$ which gives $\mathcal{C} = \mathcal{C}(y\theta; x) = (x + y\theta)b_1^{R/(1-R)}G(p)^{1/(1-R)} - x$.

Theorem 6.4 is proved in Appendix E.

Theorem 6.4 Suppose $b_1 > 0$.
  1) $(1 - R)b_1^R G$ is decreasing in b_1.
  2) $(1 - R)G$ is increasing in b_3.

Corollary 6.5 Suppose $b_1 > 0$. $\mathcal{C}$ is decreasing in δ and increasing in α.

Both these monotonicities are intuitively natural. For the monotonicity in α, since the agent only ever holds long[8] positions in the illiquid asset, we expect him to benefit

---

[8]Note that if he starts with a solvent initial portfolio, but a negative holding in the illiquid asset, then the agent makes an instantaneous transaction at time zero to make his holding positive.

from an increase in the return and hence the future price of the illiquid asset. (Note that some care is needed in making this argument precise. Part of the optimal strategy is to sometimes purchase units of the illiquid asset, and this will potentially be more costly as α increases since by Corollary 6.3, p_* is increasing in α and thus typically, the agent may expect to make purchases of the illiquid asset at a higher price.) For the monotonicity in δ, we note that increases in δ tilt consumption towards the present with the effect that the investor has less wealth at future times with which to benefit from the growth of the expected value of the risky asset.

6.2 Monotonicity with respect to transaction costs

From the discussion in Sect. 5, we have seen that transformed boundaries only depend on the round-trip transaction cost ξ. In particular, q_* and q^* are respectively strictly decreasing and increasing in ξ and hence the Merton line is included in $[q_*, q^*]$, the no-transaction region with its critical boundaries measured in prices after transaction costs. However, the purchase/sale boundaries in the original scale still depend on the individual costs of purchase and sale. Write

$$p_*(\lambda, \gamma) = \frac{q_*(\xi)}{1 + \lambda(1 - q_*(\xi))}, \qquad p^*(\lambda, \gamma) = \frac{q^*(\xi)}{1 - \gamma(1 - q^*(\xi))}$$

and recall that $\xi = \frac{\lambda + \gamma}{1 - \gamma}$. We have

$$\frac{dp_*}{d\gamma} = \frac{\partial p_*}{\partial q_*} \frac{\partial q_*}{\partial \xi} \frac{\partial \xi}{\partial \gamma} = \frac{1 + \lambda}{(1 - \gamma)^2} \frac{1 + \lambda}{(1 + \lambda(1 - q_*))^2} \frac{\partial q_*}{\partial \xi} < 0$$

so that the critical ratio of wealth in the illiquid asset to paper wealth at which the agent purchases more illiquid asset is decreasing in the transaction cost on sales. However, the dependence of the critical ratio p_* at which purchases occur on the transaction cost on purchases is ambiguous in sign; indeed,

$$\frac{dp_*}{d\lambda} = \frac{\partial p_*}{\partial \lambda} + \frac{\partial p_*}{\partial q_*} \frac{\partial q_*}{\partial \xi} \frac{\partial \xi}{\partial \lambda} = -\frac{q_*(1 - q_*)}{(1 + \lambda(1 - q_*))^2} + \frac{1}{1 - \gamma} \frac{1 + \lambda}{(1 + \lambda(1 - q_*))^2} \frac{\partial q_*}{\partial \xi}$$

is not necessarily negative, for we may have $q_* > 1$. Indeed, the Merton line can lie outside the no-transaction region $[p_*, p^*]$, and the boundaries of this region need not be monotonic in the transaction cost parameters. In the single-asset case, the location of the no-transaction region is discussed by Shreve and Soner [23], and the issues are considered further in Hobson et al. [14].

7 Conclusion

In this paper, we study the Merton investment and consumption problem under transaction costs with two risky assets in the special case where transaction costs are payable on only one of the risky assets. The presence of the second risky asset, which may be used for hedging and investment purposes, makes the problem significantly

more complicated than the single-risky-asset case, but we can extend the methods of [14] to give a complete solution. Indeed, up to evaluating an integral of a known algebraic function, we can determine exactly when the problem is well-posed, and up to solving a free boundary value problem for a first order differential equation, we can determine the boundaries of the no-transaction wedge.

At the heart of our analysis is this free boundary value problem. Although the utility maximisation problem depends on many parameters describing the agent (his risk aversion and discount rate), the market (the interest rate and the drifts, volatilities and correlations of the traded assets) and the frictions (the transaction costs on sales and purchases), the ODE depends on the risk aversion parameter and just three further parameters, and the solution we want can be specified further in terms of the round-trip transaction cost.

Choi et al. [6] and our previous work [14] give a solution to the problem in the case of a single risky asset. The major issue in [6] and [14] is to understand the solution of an ODE as it passes through a singular point. In the present paper, the problem is richer and the ODE is more complicated, but in other ways, the analysis is much simpler because although the key ODE has singularities, these can be removed.

In the paper, we have assumed a single illiquid asset and just one further risky asset, but the analysis extends immediately to the case of a single illiquid asset and several risky assets on which no transaction costs are payable, at the expense of a more complicated notation. The extension to a model with many risky assets with transaction costs payable on all of them remains a challenging open problem.

Appendix A: Continuity and smoothness of the candidate value function

Suppose there exists a solution $(n(\cdot), q_*, q^*)$ to Problem 3.2 with n being strictly positive. Define the constants $p_* = \frac{q_*}{1+\lambda(1-q_*)}$ and $p^* = \frac{q^*}{1-\gamma(1-q^*)}$ and the functions $N(q) = \mathrm{sgn}(1-q)n(q)^{-R}|1-q|^{R-1}$, $W = N^{-1}$ (which is the inverse of N) and $w(h) = (1-R)hW(h)$. We should like to construct the candidate value function G^C based on the definition $G^C(p) = \mathrm{sgn}(1-p)|1-p|^{1-R}h(p)$, where h solves $\frac{dh}{dp} = \frac{w(h)}{p(1-p)}$. The main subtlety is that $\frac{w(h)}{p(1-p)}$ is not well defined at $p=1$. Nonetheless, the definition of G^C at $p=1$ can be understood in a limiting sense. To this end, we distinguish two different cases based on whether $q_* - 1$ and $q^* - 1$ have the same sign or not, or equivalently whether the no-transaction wedge, plotted in $(x, y\theta)$ space, includes the vertical axis $\{x=0\}$ (corresponding to $p=1$).

Proposition A.1 (i) *Suppose* $1 \notin [p_*, p^*]$. *Define* $h(p)$ *via*

$$\int_{N(q_*)}^{h(p)} \frac{du}{w(u)} = \int_{p_*}^{p} \frac{du}{u(1-u)} \tag{A.1}$$

on $\{p_* \leqslant p \leqslant p^*\}$. *Then* (A.1) *is equivalent to*

$$\int_{h(p)}^{N(q^*)} \frac{du}{w(u)} = \int_{p}^{p^*} \frac{du}{u(1-u)}, \tag{A.2}$$

and (A.2) *is an alternative definition of* $h(p)$.

   Let

$$G^C(p) = \begin{cases} n(q_*)^{-R} (1+\lambda p)^{1-R}, & p \in [-\frac{1}{\lambda}, p_*), \\ \text{sgn}(1-p)|1-p|^{1-R} h(p), & p \in [p_*, p^*], \\ n(q^*)^{-R} (1-\gamma p)^{1-R}, & p \in (p^*, \frac{1}{\gamma}]. \end{cases}$$

Then G^C *is a* C^2 *function on* $(-\frac{1}{\lambda}, \frac{1}{\gamma})$. *Moreover,* $\frac{(x+y\theta)^{1-R}}{1-R} G^C(\frac{y\theta}{x+y\theta})$ *is strictly increasing and strictly concave in* x.

   (ii) *Suppose* $1 \in [p_*, p^*]$. *Define* $h(p)$ *via*

$$\begin{cases} \int_{N(q_*)}^{h(p)} \frac{du}{w(u)} = \int_{p_*}^{p} \frac{du}{u(1-u)}, & p_* < p < 1, \\ \int_{h(p)}^{N(q^*)} \frac{du}{w(u)} = \int_{p}^{p^*} \frac{du}{u(1-u)}, & 1 < p < p^*. \end{cases}$$

Let

$$G^C(p) = \begin{cases} n(q_*)^{-R} (1+\lambda p)^{1-R}, & p \in [-\frac{1}{\lambda}, p_*), \\ \text{sgn}(1-p)|1-p|^{1-R} h(p), & p \in [p_*, p^*] \setminus \{1\}, \\ n(1)^{-R} e^{-(1-R)a}, & p = 1, \\ n(q^*)^{-R} (1-\gamma p)^{1-R}, & p \in (p^*, \frac{1}{\gamma}] \end{cases}$$

with $a := -\int_{q_*}^{1} \frac{R}{q(1-R)} \frac{O(q,n(q))}{n(q)} dq - \ln(1+\lambda)$. *Then* $|a| \leqslant \ln(1+\xi)$, *and* G^C *is a* C^2 *function on* $(-\frac{1}{\lambda}, \frac{1}{\gamma})$. *Moreover,* $\frac{(x+y\theta)^{1-R}}{1-R} G^C(\frac{y\theta}{x+y\theta})$ *is strictly increasing and strictly concave in* x.

Proof (i) We have

$$\int_{N(q_*)}^{N(q^*)} \frac{du}{w(u)} - \int_{p_*}^{p^*} \frac{du}{u(1-u)}$$

$$= \int_{q_*}^{q^*} \left(\frac{N'(u)}{(1-R)u N(u)} - \frac{1}{u(1-u)} \right) du + \int_{q_*}^{q^*} \frac{du}{u(1-u)} - \int_{p_*}^{p^*} \frac{du}{u(1-u)}$$

$$= \int_{q_*}^{q^*} \left(-\frac{R}{u(1-R)} \frac{O(u,n(u))}{n(u)} \right) du - \ln(1+\xi) = 0,$$

using (3.8), and this establishes the equivalence of (A.1) and (A.2).

Suppose we have a solution $(n(\cdot), q_*, q^*)$ to (3.9) with n being strictly positive. Let $G^C = G^C(p)$ be defined as at the start of this section. For notational convenience (and to allow us to write derivatives as superscripts), write G as shorthand for G^C.

First we check that G is C^2. Outside the no-transaction interval, this is immediate from the definition, and on (p_*, p^*), it follows from the fact that n and n' are continuous. This property is inherited by the pair (w, w') and then on integration by the trio (h, h', h'') and finally (G, G', G'').

It remains to check the continuity of G, G' and G'' at p_* and p^*. We prove the continuity at p_*; the proofs at p^* are similar. Using $\frac{1-q_*}{1-p_*} = \frac{1}{1+\lambda p_*}$ for the penultimate equivalence, we have

$$
\begin{aligned}
G(p_*+) &= \text{sgn}(1 - p_*)|1 - p_*|^{1-R} h(p_*) \\
&= \text{sgn}(1 - p_*)|1 - p_*|^{1-R} \text{sgn}(1 - q_*) n(q_*)^{-R}|1 - q_*|^{R-1} \\
&= n(q_*)^{-R}(1 + \lambda p_*)^{1-R} = G(p_*-).
\end{aligned}
$$

By tracing the definitions of h, w and W, it can be shown that

$$
G(p) - \frac{pG'(p)}{1 - R} = |1 - p|^{-R} h\big(1 - W(h)\big) \tag{A.3}
$$

and

$$
\begin{aligned}
&p^2 G''(p) + 2RpG'(p) - R(1 - R)G(p) \\
&= \text{sgn}(1 - p)|1 - p|^{-(1+R)}\big(w(h)w'(h) + (2R - 1)w(h) - R(1 - R)h\big). \tag{A.4}
\end{aligned}
$$

Then continuity of G' at p_* immediately follows from (A.3), where

$$
G(p_*+) - \frac{p_* G'(p_*+)}{1 - R} = \frac{h_*(1 - W(h_*))}{|1 - p_*|^R} = G(p_*-) - \frac{p_* G'(p_*-)}{1 - R}.
$$

Continuity of G'' at p_* now follows from (A.4) and continuity of G and G'.

Now we argue that $\frac{(x+y\theta)^{1-R}}{1-R} G(\frac{y\theta}{x+y\theta})$ is strictly increasing and strictly concave in x. Outside $[p_*, p^*]$, this is immediate from the definition. On $[p_*, p^*]$, the increasing property follows if $G(p) - \frac{pG'(p)}{1-R} > 0$. But this is trivial since

$$
G(p) - \frac{pG'(p)}{1 - R} = \frac{h(1 - W(h))}{|1 - p|^R} = \frac{N(q)(1 - q)}{|1 - p|^R} = \left|\frac{1 - q}{1 - p}\right|^R n(q)^{-R} > 0.
$$

Meanwhile, $\frac{(x+y\theta)^{1-R}}{1-R} G(\frac{y\theta}{x+y\theta})$ being concave on $[p_*, p^*]$ is equivalent to the condition that $qw'(N(q)) + (2R - 1)q - R < 0$. But this follows from our choice of root in (3.7).

(ii) Note that the integrand of $\int_{q_*}^{q^*} \frac{R}{q(1-R)} \frac{O(q,n(q))}{n(q)} \, dq$ is everywhere negative and therefore we have existence of $\int_{q_*}^{1} (-\frac{R}{q(1-R)} \frac{O(q,n(q))}{n(q)}) \, dq$ in $[0, \ln(1 + \xi)]$. Hence we can establish $-\ln(1 + \xi) \leqslant a \leqslant \ln(1 + \xi)$.

For $p \neq 1$, the C^2 smoothness of $G = G^C$ follows as in (i). We focus on the case $p = 1$. Suppose first that $p_* < 1 < p^*$. Continuity of G and G' at $p = 1$ can be established if we can show that both

$$\lim_{p \to 1} \frac{1}{G(p)} \left(G(p) - \frac{pG'(p)}{1-R} \right)^{1-1/R} = n(1) \tag{A.5}$$

and

$$\lim_{p \to 1} \frac{pG'(p)}{(1-R)G(p)} = 1 - e^a. \tag{A.6}$$

Substituting (A.6) into (A.5), we recover the given value of $G(1)$.

Using (A.3) and the equivalence of $p \to 1$ and $q \to 1$, we have

$$\frac{1}{G(p)} \left(G(p) - \frac{pG'(p)}{1-R} \right)^{1-1/R} = |h|^{-1/R} |1 - W(h)|^{1-1/R} = n(q) \longrightarrow n(1)$$

and (A.5) holds. For (A.6), we have

$$\frac{1 - W(h(p))}{1-p} = \frac{(1-R)h(p) - p(1-p)h'(p)}{(1-R)(1-p)h(p)} = 1 - \frac{pG'(p)}{(1-R)G(p)}.$$

Suppose $p < 1$. Then using the definition of $h(p)$,

$$0 = \int_{N(q_*)}^{h(p)} \frac{du}{w(u)} - \int_{p_*}^{p} \frac{du}{u(1-u)} = \int_{q_*}^{W(h(p))} \frac{N'(q)\,dq}{(1-R)qN(q)} - \int_{p_*}^{p} \frac{du}{u(1-u)}$$

$$= \int_{q_*}^{W(h(p))} \left(\frac{N'(q)}{(1-R)qN(q)} - \frac{1}{q(1-q)} \right) dq$$

$$+ \int_{q_*}^{W(h(p))} \frac{dq}{q(1-q)} - \int_{p_*}^{p} \frac{du}{u(1-u)}$$

$$= \int_{q_*}^{W(h(p))} \left(-\frac{R}{u(1-R)} \frac{O(u, n(u))}{n(u)} \right) du - \int_{p_*}^{q_*} \frac{du}{u(1-u)} - \int_{W(h(p))}^{p} \frac{dq}{q(1-q)}$$

$$= \int_{q_*}^{W(h(p))} \left(-\frac{R}{u(1-R)} \frac{O(u, n(u))}{n(u)} \right) du - \ln \left(\frac{p}{W(h(p))} \frac{1 - W(h(p))}{1-p} \right)$$

$$- \ln(1 + \lambda).$$

Letting $p \uparrow 1$ and using $\lim_{p \to 1} W(h(p)) = 1$, we obtain $\lim_{p \uparrow 1} \frac{1 - W(h(p))}{1-p} = e^a$. A similar calculation for $p > 1$ leads to $\lim_{p \downarrow 1} \frac{W(h(p)) - 1}{p - 1} = e^a$. Hence (A.6) holds. As a byproduct, we can establish

$$\lim_{p \to 1} G'(p) = (1-R)(1 - e^a)G(1) = (1-R)(1 - e^a)n(1)^{-R} e^{-(1-R)a}.$$

Consider now continuity of G'' at $p = 1$. We show that $\lim_{p \to 1} G''(p)$ exists. Consider $\hat{D}(p) = \frac{((1-R)G(p)-pG'(p))^2}{G(p)(p^2G''(p)+2RpG'(p)-R(1-R)G(p))}$. Then

$$
\begin{aligned}
\hat{D}(p) &= \frac{(1-R)^2 h(1-W(h))^2}{w(h)w'(h) + (2R-1)w(h) - R(1-R)h} \\
&= \frac{(1-R)(1-q)^2}{(1-R)qN(q)/N'(q) - (1-q)(R+(1-R)q)} \\
&= \frac{(1-R)(1-R-R(1-q)n'(q)/n(q))}{R(R+(1-R)q)n'(q)/n(q) - R(1-R)}
\end{aligned}
$$

and

$$
\begin{aligned}
\lim_{p \to 1} \hat{D}(p) &= \lim_{q \to 1} \frac{(1-R)(1-R-R(1-q)n'(q)/n(q))}{R((R+(1-R)q)n'(q)/n(q) - (1-R))} \\
&= \frac{(1-R)^2}{R(n'(1)/n(1) - (1-R))}.
\end{aligned} \tag{A.7}
$$

Note that $n'(1)/n(1) - (1-R) \neq 0$ since $\operatorname{sgn}(n'(1)) = -\operatorname{sgn}(1-R)$. The limit is thus always well defined and can be used to obtain an expression for $\lim_{p \to 1} G''(p)$.

Since G is C^2 and (3.5) holds for both $p < 1$ and $p > 1$, it follows that (3.5) holds at $p = 1$ also and $\frac{(x+y\theta)^{1-R}}{1-R} G(\frac{y\theta}{x+y\theta})$ is concave on $[p_*, p^*]$.

Finally, we consider the case where $p_* = 1$ or $p^* = 1$. Suppose we are in the former scenario. Then to show the continuity of G at $p_* = 1$, it is sufficient to show that $n(q_*)^{-R}(1+\lambda)^{1-R} = n(1)^{-R}e^{-(1-R)a}$. But $q_* = 1$ when $p_* = 1$ and thus $a = -\ln(1+\lambda)$. The above expression then holds immediately. The values of $G'(1)$ and $G''(1)$ can again be inferred from (A.6) and (A.7). A similar result follows in the case $p^* = 1$. $\qquad\square$

Appendix B: The candidate value function and the HJB equation

We have to verify that the candidate value function given in Proposition A.1 solves the HJB variational inequality

$$
\min\left\{ -\sup_{c>0,\pi} \mathcal{L}^{c,\pi} V^C, -\mathcal{M}V^C, -\mathcal{N}V^C \right\} = 0, \tag{B.1}
$$

where $\mathcal{L}$, $\mathcal{M}$ and $\mathcal{N}$ are the operators

$$
\mathcal{L}^{c,\pi} f := \frac{c^{1-R}}{1-R} - cf_x + \frac{\sigma^2}{2}f_{xx}\pi^2 + ((\mu-r)f_x + \sigma\eta\rho f_{xy}y)\pi
$$

$$
+ rf_x x + \alpha f_y y + \frac{\eta^2}{2}f_{yy}y^2 - \delta f,
$$

$$
\mathcal{M}f := f_\theta - (1+\lambda)yf_x, \qquad \mathcal{N}f := (1-\gamma)yf_x - f_\theta.
$$

Outside the no-transaction region, most of the inequalities in (B.1) follow from the construction of V^C or direct substitution. We provide a proof of the less trivial result that $\mathcal{M}V^C \leqslant 0$ on the no-transaction region $\{p \in [p_*, p^*]\}$. The inequality $\mathcal{N}V^C \leqslant 0$ can be proved in an identical fashion. Again writing G as shorthand for G^C, we have

$$\mathcal{M}V^C = V_\theta^C - (1+\lambda)yV_x^C = \frac{pV^C}{\theta}\left((1+\lambda p)\frac{G'(p)}{G(p)} - \lambda(1-R)\right).$$

Since $\operatorname{sgn}(V^C) = \operatorname{sgn}(1-R)$, it is necessary and sufficient to show

$$\operatorname{sgn}(1-R)\left((1+\lambda p)\frac{G'(p)}{G(p)} - \lambda(1-R)\right) \leqslant 0.$$

But $G(p) = \operatorname{sgn}(1-p)h(p)|1-p|^{1-R}$ for $p \neq 1$, and then

$$\frac{G'(p)}{G(p)} = \frac{h'(p)}{h(p)} - \frac{1-R}{1-p} = \frac{w(h)}{h(p)p(1-p)} - \frac{1-R}{1-p} = \frac{1-R}{1-p}\left(\frac{W(h)}{p} - 1\right)$$

and the required inequality becomes

$$\frac{1-W(h)}{1-p} \geqslant \frac{1}{1+\lambda p}. \tag{B.2}$$

We are going to prove (B.2) for $p \in [p_*, p^*]\setminus\{1\}$. Then $\mathcal{M}V^C \leqslant 0$ will hold at $p=1$ as well by smoothness of V^C.

By construction, $q = W(h(p))$. Since W is monotonic and h is monotonic except possibly at $p=1$, it follows that q is an increasing function of p. Then, starting from the identity $\int_{N(q_*)}^{N(q)} \frac{dh}{w(h)} = \int_{p_*}^{p} \frac{du}{u(1-u)}$ and following the substitutions leading to (3.8), we find

$$\int_{q_*}^{q}\left(-\frac{R}{u(1-R)}\frac{O(u,n(u))}{n(u)}\right)du = -\int_{q_*}^{q}\frac{dv}{v(1-v)} + \int_{p_*}^{p}\frac{du}{u(1-u)}.$$

Since the expression on the left-hand side is increasing in q, we deduce that we have $\frac{1}{q(1-q)}\frac{dq}{dp} \leqslant \frac{1}{p(1-p)}$.

Define $\chi(p) := \frac{(1+\lambda)p}{1+\lambda p}$. Then χ is a solution to the ODE $\chi'(p) = \varrho(p, \chi(p))$, where $\varrho(p,y) = \frac{y(1-y)}{p(1-p)}$. Note that $\chi(p_*) = \frac{(1+\lambda)p_*}{1+\lambda p_*} = q_* = q(p_*)$.

Suppose $p_* < p^* < 1$. Then for $p < 1$ and in turn $q = q(p) = W(h(p)) < 1$, we have $q'(p) \leqslant \varrho(p, q(p))$ and conclude that $q(p) \leqslant \chi(p)$ for $p_* \leqslant p < p^* \leqslant 1$. Then $1 - W(h(p)) = 1 - q(p) \geqslant 1 - \chi(p) = \frac{1-p}{1+\lambda p}$ which establishes (B.2). If instead $1 < p_* < p^*$, we can arrive at the same result by showing $q(p) \geqslant \chi(p)$ for $1 < p_* \leqslant p$ and in turn $\frac{dq}{dp} \geqslant \frac{q(q-1)}{p(p-1)}$.

It remains to consider the case of $p_* \leqslant 1 \leqslant p^*$. The only issue is that the comparison of derivatives of $q(p)$ and $\chi(p)$ may be not trivial at $p=1$ because of the singularity in $\varrho(p,y)$. But by direct computation, it can be found that $\chi'(1) = \frac{1}{1+\lambda}$. Meanwhile, $q'(1-) = \lim_{p\uparrow 1}\frac{1-q(p)}{1-p} = \lim_{p\uparrow 1}\frac{1-W(h(p))}{1-p} = e^a$, and similarly, we have

$q'(1+) = e^a$. Then $q'(1)$ is well defined, and moreover since $a > -\ln(1+\lambda)$, we have $q'(1) = e^a > 1/(1+\lambda) = \chi'(1)$. Together with the fact that $q(1) = 1 = \chi(1)$, we must have that $q(p)$ is an upcrossing of $\chi(p)$ at $p = 1$. From this, we conclude that $q(p) \leqslant \chi(p)$ on $\{p \in [p_*, 1)\}$ and $\chi(p) \leqslant q(p)$ on $\{p \in (1, p^*]\}$. And (B.2) then follows.

Appendix C: Proof of the main results

Proof of Theorems 4.1 and 4.3 We prove the two theorems together. Suppose we are in the well-posed cases. From the analysis in Sect. 5, there exists a solution $(n(\cdot), q_*, q^*)$ to the free boundary value problem with n being strictly positive. By the C^2 smoothness of G^C, V^C is $C^{2,2,1}$. Moreover, in Appendices A and B, we have seen that V^C is a strictly concave function in x solving the HJB variational inequality (B.1). For $R < 1$, using the supermartingale property of $M_t := \int_0^t e^{-\delta s} \frac{C_s^{1-R}}{1-R} ds + e^{-\delta t} V^C(X_t, Y_t, \Theta_t)$, we can establish $V \leq V^C$ following standard arguments. For $R > 1$, M may be only a local supermartingale, and further arguments along the lines of those in Davis and Norman [9] or Tse [25, Appendix 3.D] are needed.

To show $V^C \leqslant V$, it is sufficient to demonstrate the existence of an investment/consumption strategy which attains the value V^C. Suppose the initial value $(x, y\theta)$ is such that $\frac{y\theta}{x+y\theta} = p \in [p_*, p^*]$. Define feedback controls $C^* = (C_t^*)_{t \geqslant 0}$ and $\Pi^* = (\Pi_t^*)_{t \geqslant 0}$, where $C_t^* = C^*(X_t, Y_t, \Theta_t)$ and $\Pi_t^* = \Pi^*(X_t, Y_t, \Theta_t)$ with $C^*(x, y, \theta) := (V_x^C(x, y, \theta))^{-\frac{1}{R}}$,

$$\Pi^*(x, y, \theta) := -\frac{(\mu - r)V_x^C(x, y, \theta) + \sigma \eta \rho y V_{xy}^C(x, y, \theta)}{\sigma^2 V_{xx}^C(x, y, \theta)},$$

and where $\Theta^* = (\Theta_t^*)_{t \geqslant 0}$ is a finite-variation, local-time-type strategy of the form $\Theta_t^* = \theta + \Phi_t^* - \Psi_t^*$ which keeps (P_t) within (p_*, p^*). Our goal is to show that the process $M^* = (M_t^*)_{t \geqslant 0}$, which is defined as a version of M evolving under (C^*, Π^*, Θ^*), is a true martingale. The technical delicacy is to show that the local-martingale stochastic integrals $\int_0^t e^{-\delta s} \sigma V_x^C \Pi_s^* dB_s$ and $\int_0^t e^{-\delta s} \eta V_y^C Y_s dW_s$ are indeed true martingales. But this can be done following ideas similar to Davis and Norman [9]. Further, it can also be shown that $\lim_{t \to \infty} \mathbb{E}[e^{-\delta t} V^C(X_t^*, Y_t, \Theta_t^*)] = 0$. Then standard arguments lead to $V^C \leq V$. See Tse [25, Appendix 3.E] for a detailed proof.

Now suppose the initial value $(x, y\theta)$ is such that $p < p_*$. Then consider the strategy of purchasing $\phi = \frac{xp_* - (1-p_*)y\theta}{y(1+\lambda p_*)}$ shares at time zero so that the post-transaction proportional holding in the illiquid asset is $\frac{y(\theta+\phi)}{x+y(\theta+\phi)-y(1+\lambda)\phi} = p_*$, and thereafter following the investment/consumption strategy (C^*, Π^*, Θ^*) as in the case of $\{p \in [p_*, p^*]\}$. By its construction, $V^C(x, y, \theta) = V^C(x - y(1+\lambda)\phi, y, \theta + \phi)$, and from this we can conclude that $V^C \leqslant V$. A similar argument applies for an initial value with $p > p^*$.

Finally, we consider the conditionally well-posed case. From the discussion in Sect. 5, it is clear that as long as $\xi > \bar{\xi}$, there still exists a solution $(n(\cdot), q_*, q^*)$ to the

free boundary value problem, and thus one can show $V^C = V$ by the same argument as for the unconditionally well-posed cases. Moreover, from Lemma 5.1, we can see that $n(\cdot) \downarrow 0$ as $\xi \downarrow \overline{\xi}$, and in turn $V^C \to \infty$ from its construction. But $V \geqslant V^C$, and thus we conclude that $V \to \infty$ as $\xi \downarrow \overline{\xi}$. This shows the ill-posedness of the problem at $\xi = \overline{\xi}$, and using the monotonicity of V in ξ, this extends to any $\xi \leqslant \overline{\xi}$. □

Proof of Corollary 4.2 Note that if $R < 1$, then $m_M \leq m(1) \leq \ell(1)$ so that $\ell(1) \leq 0$ is necessary and sufficient for both $m_M < 0$ and $\ell(1) \leq 0$. Further, $\ell(1) \leqslant 0$ is equivalent to $b_3 \geqslant \frac{b_1}{1-R} + b_2 R$, and this inequality can be restated as $\alpha \geqslant \frac{1}{2}\eta^2 R + \frac{\delta}{1-R}$. But this is exactly the ill-posedness condition in the one-risky-asset case; see [14] or [6]. □

Appendix D: The first order differential equation

The goal of this section is to establish some important results regarding the functions $m(q)$, $\ell(q)$ and $O(q, n)$, which then allow us to infer the properties of the solution to the ODE $n'(q) = O(q, n(q))$ as in Sect. 5.

Let $\mathcal{S} \subseteq \{(q, n); q > 0, n \geq 0\}$ be the set

$$\mathcal{S} = \{q = 1\} \cup \left\{ q = \frac{R}{R-1} \right\} \cup \{n = 0\} \cup \{q < 1, (1 - R)n \geq (1 - R)\ell(q)\}.$$

On $(0, \infty) \times [0, \infty) \setminus \mathcal{S}$, define $F(q, n) = O(q, n)/n$. Extend the definition of F to $(0, \infty) \times [0, \infty)$ where possible by taking appropriate limits. The lemma below collects all the relevant results; it is an extended version of Lemma 3.1.

Lemma D.1 1) (a) *For $R < 1$, $\ell(q) > m(q)$ on $\{q \in (0, 1]\}$. Moreover, on $(0, \infty)$, m crosses ℓ exactly once from below at some point above* 1.

(b) *For $R > 1$, $m(q) > \ell(q)$ on $\{q \in (0, 1]\}$. Moreover, on $(0, \infty)$, m either does not cross ℓ at all, or touches ℓ exactly once in the open interval $(1, R/(R - 1))$, or crosses ℓ twice on $(1, R/(R - 1))$.*

2) *For $R > 1$, $F(q, n)$ is well defined at $q = R/(R - 1)$.*

3) *For $n > 0$ and $(1 - R)n < (1 - R)\ell(1)$, $F(1, n)$ is well defined and*

$$F(1, n) := \lim_{q \to 1} F(q, n) = -\frac{(1 - R)(n - m(1))}{\ell(1) - n}. \tag{D.1}$$

Also, for $q \leq 1$ and $R < 1$, we have $\lim_{n \uparrow \ell(q)} F(q, n) = -\infty$ (and if $R > 1$, $\lim_{n \downarrow \ell(q)} F(q, n) = +\infty$). For $q > 1$ and $R < 1$ (and $1 < q < \frac{R}{R-1}$ for $R > 1$), we have that $F(q, \ell(q)) := \lim_{n \to \ell(q)} F(q, n)$ satisfies

$$F(q, \ell(q)) = -\frac{1 - R}{R(1 - q)} \left(\frac{q((1 - R)q + R)}{((1 - R)q + R)^2 + (b_2 - 1)R^2} - 1 \right). \tag{D.2}$$

4) *$F(q, n) = 0$ if and only if $n = m(q)$. Moreover,*
(a) *for $R < 1$:*

(i) *On $\{0 < q < 1\}$, $F(q, n) < 0$ for $m(q) < n < \ell(q)$ and $F(q, n) > 0$ for $n < m(q)$ or $n > \ell(q)$.*

(ii) *At $q = 1$, $F(1, n) < 0$ for $m(1) < n < \ell(1)$ and $F(1, n) > 0$ for $n < m(1)$; $F(1, n)$ is not well defined for $n \geqslant \ell(1)$.*

(iii) *On $\{q > 1\}$, $F(q, n) < 0$ for $n > m(q)$ and $F(q, n) > 0$ for $n < m(q)$.*

(b) *for $R > 1$:*

(i) *On $\{0 < q < 1\}$, $F(q, n) > 0$ for $\ell(q) < n < m(q)$ and $F(q, n) < 0$ for $n < \ell(q)$ or $n > m(q)$.*

(ii) *At $q = 1$, $F(1, n) > 0$ for $\ell(1) < n < m(1)$ and $F(1, n) < 0$ for $n > m(1)$; $F(1, n)$ is not well defined for $n \leqslant \ell(1)$.*

(iii) *On $\{1 < q \leqslant R/(R-1)\}$, $F(q, n) < 0$ for $n > m(q)$ and $F(q, n) > 0$ for $n < m(q)$.*

(iv) *On $\{q > R/(R-1)\}$, $F(q, n) < 0$ for $m(q) < n < \ell(q)$ and $F(q, n) > 0$ for $n > \ell(q)$ or $n < m(q)$.*

Before we proceed, we introduce some additional notation:

$$v(q, n) = \varphi(q, n) - \mathrm{sgn}(1 - R)\sqrt{\varphi(q, n)^2 + E(q)^2},$$

$$D(q, n) = 2\big((1 - R)q + R\big)\big(n - m(q)\big) - q\Big(v(q, n) - v\big(q, m(q)\big)\Big),$$

$$A(q, n) = D(q, n)$$
$$+ \big(\ell(q) - n\big)\Bigg(2\big((1 - R)q + R\big) - q\Big(1 - \mathrm{sgn}(1 - R)\frac{\varphi}{\sqrt{\varphi^2 + E^2}}\Big)\Bigg),$$
(D.3)

where m, ℓ, φ and E are defined in Sect. 3. We begin with a useful lemma, whose proof is a lengthy exercise in algebra and is omitted.

Lemma D.2 *$O(q, n)$ has the alternative expression*

$$O(q, n) = -\frac{(1 - R)n D(q, n)}{2R(1 - q)((1 - R)q + R)(\ell(q) - n)}.$$
(D.4)

Proof of Lemma D.1 1) Observe that

$$\ell(q) - m(q) = \frac{(1 - R)q}{((1 - R)q + R)}P(q),$$

where $P(q) := Rb_2 + (1 - 2R)q - (1 - R)q^2$. Hence the crossing points of $\ell(q)$ and $m(q)$ away from $q = 0$ are given by the roots of $P(q) = 0$ if such roots exist. The desired results can be established easily by studying the quadratic function $P(q)$.

2) The behaviour at $q = -R/(1 - R)$ is only relevant for $R > 1$; so we write this as $q = R/(R - 1)$. Note that ℓ explodes at $q = \frac{R}{R-1}$. It is sufficient to check that the denominator of $O(q, n)$ is not equal to zero at $q = R/(R - 1)$. Direct calculation

gives $((1 - R)q + R)(\ell(q) - n)|_{q=\frac{R}{R-1}} = -(b_2 - 1)R^2$ and hence

$$2R(1-q)\big((1-R)q + R\big)(\ell(q) - n)\big|_{q=\frac{R}{R-1}} = \frac{2R^3(b_2 - 1)}{(R-1)} \neq 0.$$

3) Both the numerator and denominator of $F(q, n)$ are zero at $q = 1$. L'Hôpital's rule can be applied to calculate $\lim_{q \to 1} \frac{D(q,n)}{1-q}$ to deduce the expression in (D.1) after some algebra.

Now consider $\lim_{n \to \ell(q)} F(q, n)$. Suppose first $0 < q < 1$. Then

$$D\big(q, \ell(q)\big) = 2(1 - R)q(1 - q)\left(\big((1 - R)q + R\big) + \frac{R^2(b_2 - 1)}{(1 - R)q + R}\right),$$

which is non-zero and has $\text{sgn}(D(q, \ell(q))) = \text{sgn}(1 - R)$. It follows that for $q < 1$ and $R < 1$, $\lim_{n \uparrow \ell(q)} F(q, n) = -\infty$ and for $q < 1$ and $R > 1$, $\lim_{n \downarrow \ell(q)} F(q, n) = +\infty$.

Now suppose $q > 1$, and if $R > 1$ that $(1 - R)q + R > 0$. Then direct evaluation gives $D(q, \ell(q)) = 0$.[9] In order to determine the value of $F(q, \ell(q))$ via L'Hôpital's rule, we need to compute

$$\frac{\partial D}{\partial n} = 2\big((1 - R)q + R\big) - q\frac{\partial v}{\partial n} = 2\big((1 - R)q + R\big) - q\left(1 - \frac{\text{sgn}(1 - R)\varphi}{\sqrt{\varphi^2 + E^2}}\right), \quad \text{(D.5)}$$

and (D.2) follows immediately.

4) We prove the results for $R < 1$. The results for $R > 1$ can be obtained similarly, the only issue being that there is an extra case which arises when $(1 - R)q + R$ changes sign.

Note that for fixed q, the ordering of $m(q)$ and $\ell(q)$ is given by 1). The monotonicity of F in n for $q = 1$ can be obtained from (D.1).

If $0 < q < 1$, then since

$$2\big((1 - R)q + R\big) - q\left(1 - \frac{\text{sgn}(1 - R)\varphi}{\sqrt{\varphi^2 + E^2}}\right) > 2\big((1 - R)q + R\big) - 2q$$

$$= 2R(1 - q) > 0,$$

we conclude from (D.5) that $D(q, n)$ is increasing in n. Since $D(q, m(q)) = 0$, it follows that $D(q, n) > 0$ for $n > m(q)$ and $D(q, n) < 0$ for $n < m(q)$. Hence, $F(q, n) = 0$ if and only if $n = m(q)$, and we have

$$\text{sgn}\big(F(q, n)\big) = -\text{sgn}\left(\frac{D(q, n)}{(1 - q)((1 - R)q + R)(\ell(q) - n)}\right)$$

$$= \text{sgn}\big((n - m(q))(n - \ell(q))\big).$$

This gives the desired sign properties of $F(q, n)$ on the range $\{0 < q < 1\}$.

---

[9]More generally, $D(q, \ell(q)) = 0$ if and only if $(1 - q)((1 - R)q + R) < 0$. This can be verified with special care taken to choose the appropriate square root arising in $v(q, n)$.

Now consider the case $q > 1$ under which we have $D(q, \ell(q)) = 0$. We can compute the second derivative of D with respect to n as

$$\frac{\partial^2 D}{\partial n^2} = \text{sgn}(1 - R) q \frac{E^2}{(E^2 + \varphi^2)^{3/2}}$$

and as $R < 1$, $D(q, n)$ is convex in n. Since $D(q, m(q)) = D(q, \ell(q)) = 0$, it follows that for $q > 1$, we must have $D(q, n) < 0$ when n lies between $m(q)$ and $\ell(q)$, and $D(q, n) > 0$ otherwise. Thus $\text{sgn}(D(q, n)) = \text{sgn}((n - m(q))(n - \ell(q)))$. Then

$$\text{sgn}\left(F(q, n)\right) = \text{sgn}\left(\frac{D(q, n)}{\ell(q) - n}\right) = -\text{sgn}\left(n - m(q)\right).$$

Finally, note that $F(q, n)$ can be zero only if $n = m(q)$ or $n = \ell(q)$. But for $q > 1$, the limiting expression at $n = \ell(q)$ is given by 3). Hence $F(q, n) = 0$ if and only if $n = m(q)$. $\qquad \square$

The following lemma on further properties of F is key in the proofs of the monotonicity property of Σ and in the results on comparative statics.

Lemma D.3 *For $q \in (0, 1]$ and $(1 - R)m(q) < (1 - R)n < (1 - R)\ell(q)$, and for $q > 1$ and $(1 - R)m(q) < (1 - R)n$, we have $\frac{\partial}{\partial n} F(q, n) \leqslant 0$.*

Proof Direct computation gives

$$\left(\ell(q) - n\right)^2 \frac{\partial}{\partial n}\left(\frac{D(q, n)}{\ell(q) - n}\right) = \left(\ell(q) - n\right)\frac{\partial D}{\partial n} + D(q, n) = A(q, n),$$

where A was defined in (D.3), and in turn $\frac{\partial}{\partial n} A(q, n) = \text{sgn}(1 - R)\frac{E(q)^2 q(\ell(q) - n)}{(\varphi^2 + E(q)^2)^{3/2}}$. Hence for $q > 0$ and $R < 1$, $A(q, n)$ is increasing in n for $n < \ell(q)$ and decreasing in n for $n > \ell(q)$. If $R > 1$, then $A(q, n)$ is decreasing in n for $n < \ell(q)$ and increasing in n for $n > \ell(q)$.

Now we consider the limiting value of $A(q, n)$ as $n \to \pm\infty$. We can compute

$$\lim_{(1-R)n \to +\infty} A(q, n) = 2(1 - R)\left((1 - R)q + R\right)q(1 - q),$$

$$\lim_{(1-R)n \to -\infty} A(q, n) = \frac{2R^2(1 - R)(b_2 - 1)q(1 - q)}{(1 - R)q + R}$$

after some algebra.[10]

Suppose $R < 1$. For $0 < q < 1$, we have that $A(q, n)$ is increasing in n for $n < \ell(q)$ and decreasing in n for $n > \ell(q)$. Since on this range of q, we have

---

[10]The case of $(1 - R)n \to -\infty$ might appear to be more difficult since $v(q, n)$ does not converge, but computation can be facilitated by a useful observation that $\varphi(q, n) - n$ is independent of n.

$$\lim_{n \to +\infty} A(q, n) = 2(1 - R)((1 - R)q + R)q(1 - q) > 0,$$

$$\lim_{n \to -\infty} A(q, n) = \frac{2R^2(1 - R)(b_2 - 1)q(1 - q)}{(1 - R)q + R} > 0,$$

we conclude that $A(q, n) > 0$ for all n. If $q > 1$, then $A(q, \ell(q)) = D(q, \ell(q)) = 0$. But $A(q, n)$ attains its maximum at $n = \ell(q)$; hence we have $A(q, n) \leqslant 0$ for $q > 1$. Putting the cases together, $(1 - q)A(q, n) \geq 0$ and $\frac{\partial F}{\partial n} \leq 0$. In the special case $q = 1$, the result follows from (D.1).

Similar arguments can be adopted if $R > 1$, with extra care taken towards potential changes of sign at $q = \frac{R}{R-1}$. $\square$

Proof of Lemma 5.1 We can deduce that $\Sigma(u)$ is decreasing by using the monotonicity of $n_u(\cdot)$ in u and in turn the monotonicity of $F(q, n) = \frac{O(q,n)}{n}$ in n. When q_M is in the domain of Σ, then $\lim_{u \uparrow q_M} \Sigma(u) = 0$ can be shown easily by using the fact that q_M is a turning point of $m(q)$.

We next show that if $R < 1$ and $b_1 > 0$, then $\lim_{u \downarrow 0} \Sigma(u) = +\infty$. Suppose $R < 1$ and consider the quadratic function

$$H(x) = (1 - R)b_1(m'(0) - x) - R(l'(0) - x)x;$$

then trivially $H(m'(0)) > 0 > H(0)$. Now choose a constant k such that we have $m'(0) < k < \alpha < 0$, where α is the negative root of $H(x) = 0$. Then $H(k) > 0$ and equivalently $k < \frac{(1-R)b_1(m'(0)-k)}{R(l'(0)-k)}$. Let $b(q) = b_1 + kq$. It is clear from the definition of D that $D(0, b_1) = 0$. Thus by direct computation, we have

$$\frac{d}{dq} D(q, b_1 + kq)\Big|_{q=0} = -2Rm'(0) + 2Rk$$

and

$$\lim_{q \downarrow 0} O(q, b(q)) = -\frac{b_1(1 - R)\frac{d}{dq} D(q, b_1 + kq)|_{q=0}}{2R^2(\ell'(0) - k)} = \frac{(1 - R)b_1(m'(0) - k)}{R(l'(0) - k)}.$$

Then for all $\epsilon > 0$, there exists $K_\epsilon \in (0, 1)$ such that for $q < K_\epsilon$, we have $O(q, b(q)) > \frac{(1-R)b_1(m'(0)-k)}{R(l'(0)-k)} - \epsilon$. Pick ϵ such that $0 < \epsilon < \frac{(1-R)b_1(m'(0)-k)}{R(l'(0)-k)} - k$. Then we have $O(q, b(q)) > k$ on $\{0 < q < K_\epsilon\}$ and solutions to $n' = O(q, n)$ cross $b(q)$ from below. Let $\psi_u = \inf(q \geqslant u : n_u(q) > b(q))$. Then for $u < q < K_\epsilon \wedge \psi_u$, $n'_u(q) = O(q, n_u(q)) > O(q, b(q)) > k$. Moreover, there also exists K_m such that $m'(q) < \frac{1}{2}(m'(0) + k)$ for $q < K_m$. Hence on $\{u < q < K_\epsilon \wedge \psi_u \wedge K_m\}$, we have $n'_u(q) - m'(q) > k - \frac{1}{2}(m'(0) + k) = \frac{1}{2}(k - m'(0)) =: \widehat{k} > 0$ and then $n_u(q) - m(q) > \widehat{k}(q - u)$. On the other hand, for $\psi_u < q < K_\epsilon \wedge K_m$, $m(q) < b_1 + \frac{q}{2}(m'(0) + k)$ and hence

$$n_u(q) - m(q) > (b_1 + kq) - \left(b_1 + \frac{q}{2}(m'(0) + k)\right) = \widehat{k}q > \widehat{k}(q - u).$$

We conclude that $n_u(q) - m(q) > \widehat{k}(q - u)$ for $u < q < Q := K_\epsilon \wedge K_m$.

Hence, using ([D.4](#)) and L'Hôpital's rule,

$$\ln\left(1 + \Sigma(u)\right)$$

$$= \int_u^{\xi(u)} \left(-\frac{R}{(1-R)q} \frac{O(q, n_u(q))}{n_u(q)}\right) dq$$

$$> \int_u^Q \frac{2((1-R)q + R)(n_u(q) - m(q)) - q(v(q, n_u(q)) - v(q, m(q)))}{2q(1-q)((1-R)q + R)(\ell(q) - n_u(q))} dq.$$

Using $n_u(q) - m(q) > \widehat{k}(q - u)$, it is not hard to see that the above integrand is bounded below by $\frac{C(q-u)}{q^2}$ on $\{u < q < \widehat{Q} < Q\}$, where C and $\widehat{Q}$ are some constants independent of u. Then we deduce that

$$\ln\left(1 + \Sigma(u)\right) > \int_u^{\widehat{Q}} \frac{C(q-u)}{q^2} dq = C\left(\ln \frac{\widehat{Q}}{u} + \frac{u}{\widehat{Q}} - 1\right).$$

Letting $u \downarrow 0$, we conclude that $\Sigma(u) \to \infty$.

Now suppose $R < 1$, $b_1 > 0$, $m_M < 0$ and $\ell(1) > 0$. The smallest nonnegative solution n_u is for $u = p_-$ and then $\zeta(p_-) = p_+$ and $n_{p_-}(q) = 0$ for $q \in [p_-, p_+]$. We have $\Sigma(p_-) = \overline{\xi}$.

Now suppose $R > 1$. Most of the analysis is the same as for $R < 1$. However, when $b_1 < 0$ and $m_M > 0$ (equivalently, $b_1 > -\frac{(R-1)}{4R}b_3^2$), we can only consider solutions n_u for $u \in [p_-, q_M]$. The new case arises when $\max_{0 \le q \le 1} \ell(q) > 0$. Then, since ℓ is a lower bound for n, we have for any $u > p_-$ that $n_u(q) > \ell(q)$. In particular, $\lim_{u \downarrow p_-} n_u(q) > 0$ for $q > p^\ell$. Moreover, by considering the behaviour of F near p^ℓ, we find that $\int^{p^\ell} F(q, 0) dq$ is infinite and $\lim_{u \downarrow p_-} \Sigma(u) = \infty$. $\square$

Appendix E: Comparative statics

Proof of Proposition [6.1](#) We only prove Part 2) here. The proof of Part 1) is easier and can be done using the transformation $\overline{m}(q) = m(q) - b_1$, $\overline{n}(q) = n(q) - b_1$ and $\overline{\ell}(q) = \ell(q) - b_1$.

Set $a(q) = n(q) - m(q)$. Then the original free boundary value problem becomes to solve $a'(q) = \underline{O}(q, a(q))$ subject to $a(q_*) = a(q^*) = 0$, where

$$\underline{O}(q, a) = -\frac{(1-R)(a + m(q))D(q, a + m(q))}{2R(1-q)((1-R)q + R)(\ell(q) - m(q) - a)}$$
$$- 2R(1-R)q + b_3(1-R).$$

Observe that $\ell(q) - m(q) = (1-R)q(1-q) + (b_2 - 1)R(1-R)\frac{q}{(1-R)q+R}$ does not depend on b_3. Further, $\varphi(q, a + m(q)) = a + R(1-R)((1-q)^2 - (b_2 - 1))$ and $v(q, m(q)) = -2R(1-R)(b_2 - 1)$ are both independent of b_3. Hence

$$D\big(q, a + m(q)\big) = 2\big((1-R)q + R\big)a - q\left(v\big(q, a + m(q)\big) - v\big(q, m(q)\big)\right)$$

and

$$\frac{O\big(q, a + m(q)\big)}{a + m(q)} = -\frac{(1 - R)D(q, a + m(q))}{2R(1 - q)((1 - R)q + R)(\ell(q) - m(q) - a)}$$

are independent of b_3. Recall we are assuming $R < 1$. Then $O(q, n) \leq 0$ over the relevant range and

$$\frac{\partial \underline{O}}{\partial b_3}(q, a) = -\frac{(1 - R)D(q, a + m(q))}{2R(1 - q)((1 - R)q + R)(\ell(q) - m(q) - a)} \frac{\partial m}{\partial b_3} + 1 - R$$

$$= -\frac{O(q, a + m(q))(1 - R)q}{a + m(q)} + 1 - R > 0.$$

Let $\widehat{b}_1 > \widetilde{b}_1$ be two positive values of b_1. Define $\widehat{a}_u$ and $\widetilde{a}_u$ to be the solutions to the initial value problem $a'(q) = \underline{O}(q, a(q))$ with $a(u) = 0$ under the parameters $\widehat{b}_1$ and $\widetilde{b}_1$, respectively. We extend this notation to $\underline{O}$, ζ, Σ and (q_*, q^*).

We have $\overline{O}$ being increasing in b_3. Then $\widehat{a}_u$ cannot upcross $\widetilde{a}_u$, and since

$$\widehat{a}'_u(u) = \widehat{\underline{O}}(u, 0) - m'(u) > \widetilde{\underline{O}}(u, 0) - m'(u) = \widetilde{a}'_u(u),$$

it must be the case that $\widehat{a}_u(q) > \widetilde{a}_u(q)$ at least up to $q = \widehat{\zeta}(u) \wedge \widetilde{\zeta}(u)$. From this we conclude that $\widehat{\zeta}(u) > \widetilde{\zeta}(u)$. Hence, using the fact that $\frac{O(q, a + m(q))}{a + m(q)}$ does not depend on b_3,

$$\ln\big(1 + \widehat{\Sigma}(u)\big) = \int_u^{\widehat{\zeta}(u)} \left(-\frac{R}{q(1 - R)} \frac{O(q, \widehat{n}_u(q))}{\widehat{n}_u(q)}\right) dq$$

$$= \int_u^{\widehat{\zeta}(u)} \left(-\frac{R}{q(1 - R)} \frac{O(q, \widehat{a}_u(q) + m(q))}{\widehat{a}_u(q) + m(q)}\right) dq$$

$$> \int_u^{\widetilde{\zeta}(u)} \left(-\frac{R}{q(1 - R)} \frac{O(q, \widetilde{a}_u(q) + m(q))}{\widetilde{a}_u(q) + m(q)}\right) dq$$

$$= \ln\big(1 + \widetilde{\Sigma}(u)\big),$$

where we use the monotonicity of $\zeta(u)$ and the property that $\frac{O(q, n)}{n}$ is decreasing in n and hence $\frac{O(q, a + m(q))}{a + m(q)}$ is decreasing in a. So we get $\widehat{q}_* = \widehat{\Sigma}^{-1}(\xi) > \widetilde{\Sigma}^{-1}(\xi) = \widetilde{q}_*$. The monotonicity property of the sale boundary can be proved in a similar fashion by parametrising the family of solutions with their right boundary points. $\qquad\square$

Proof of Theorem 6.4 1) The proof is easier under a different parametrisation in which we exploit the fact that $b_1 > 0$ to write

$$V(x, y, \theta) = \left(\frac{b_1}{b_4 R}\right)^R \frac{(x + y\theta)^{1 - R}}{1 - R} \tilde{G}\left(\frac{y\theta}{x + y\theta}\right).$$

Then it is sufficient to show that $(1 - R)\tilde{G}$ is decreasing in b_1. The same transformations can be applied to $\tilde{G}$, and then it is possible to show that $\tilde{n}$ defined from $\tilde{G}$ is such

that $(1 - R)\tilde{n}$ is increasing in b_1. From this, it is possible to deduce that $(1 - R)\tilde{G}$ is decreasing in b_1; see Tse [25, Appendix 3.G] for more details.

2) Now consider the monotonicity in b_3. For $R < 1$, a similar argument to the above can be applied, but for $R > 1$, we cannot use this argument. However, the monotonicity of the value function in b_3, and hence the monotonicity of $\mathcal{C}$, can be proved by a comparison argument. The value function only depends on the parameters through R and the auxiliary parameters; so when comparing two models which differ only in b_3, we may equivalently compare two models which differ only in α.

Consider a pair of models, the only difference being that in the first model Y has drift $\tilde{\alpha}$, whereas in the second model Y has drift $\hat{\alpha} > \tilde{\alpha}$. Write $\epsilon = \hat{\alpha} - \tilde{\alpha} > 0$. Suppose that the parameters are such that the Standing Assumption 2.2 holds in the first model; then necessarily, the Standing Assumption 2.2 holds in the second model as well. Let $(\tilde{Y}_t, \hat{Y}_t) = (ye^{\eta W_t + (\tilde{\alpha} - \frac{\eta^2}{2})t}, ye^{\eta W_t + (\hat{\alpha} - \frac{\eta^2}{2})t})$ so that $\hat{Y}_t = e^{\epsilon t}\tilde{Y}_t$. Let $(\tilde{C}, \tilde{\Pi}, \tilde{\Theta} = \theta + \tilde{\Phi} - \tilde{\Psi})$ be an admissible strategy for an agent in the first model. Suppose $\tilde{\Theta}$ is nonnegative, and note that the optimal strategy has this property even if the initial endowment in the illiquid asset is negative, since in that case there is an initial transaction into the no-transaction wedge which is contained in the half-plane $\{\theta \geq 0\}$. We may assume that we start in the no-transaction region. Then $\tilde{X}_0 = x$ and $\tilde{X} = (\tilde{X}_t)_{t \geq 0}$ solves

$$d\tilde{X}_t = r(\tilde{X}_t - \tilde{\Pi}_t)\,dt + \frac{\tilde{\Pi}_t}{S_t}dS_t - \tilde{C}_t\,dt - \tilde{Y}_t(1 + \lambda)d\tilde{\Phi}_t + \tilde{Y}_t(1 - \gamma)d\tilde{\Psi}_t.$$

Set

$$\hat{\Pi}_t = \tilde{\Pi}_t, \qquad \hat{\Theta}_0 = \theta, \qquad \hat{C}_t = \tilde{C}_t + (1 - \gamma)\epsilon\tilde{\Theta}_t\tilde{Y}_t,$$

$$\hat{\Phi}_t = \int_0^t e^{-\epsilon s}d\tilde{\Phi}_s, \qquad \hat{\Psi}_t = \int_0^t e^{-\epsilon s}(d\tilde{\Psi}_s + \epsilon\tilde{\Theta}_s\,ds).$$

Note that if $\hat{\Theta}_t = \theta + \hat{\Phi}_t - \hat{\Psi}_t$, then $\hat{\Theta}_t = \tilde{\Theta}_t e^{-\epsilon t}$.

If $\hat{X}_0 = x = \tilde{X}_0$, it can be shown by simple computations that $\hat{X}$ solves the same equation as $\tilde{X}$ and $\hat{X}_t = \tilde{X}_t \geq 0$. Then for any admissible strategy in the first model for which $(\Theta_t)_{t \geq 0}$ is positive, including the optimal strategy in this model, there is a corresponding admissible strategy in the second model with strictly larger consumption at all future times. Hence the value function is strictly greater in the second model. □

References

1. Akian, M., Menaldi, J.L., Sulem, A.: Multi-asset portfolio selection problem with transaction costs. Math. Comput. Simul. **38**, 163–172 (1995)
2. Bichuch, M., Guasoni, P.: Investing with liquid and illiquid assets. Math. Finance **28**, 119–152 (2018)
3. Bichuch, M., Shreve, S.: Utility maximization trading two futures with transaction costs. SIAM J. Financ. Math. **4**, 26–85 (2013)
4. Chen, X., Dai, M.: Characterization of optimal strategy for multiasset investment and consumption with transaction costs. SIAM J. Financ. Math. **4**, 857–883 (2013)

5. Choi, J.H.: Optimal investment and consumption with liquid and illiquid assets. Math. Finance (2019, forthcoming). Available online at https://arxiv.org/abs/1602.06998
6. Choi, J.H., Sîrbu, M., Žitković, G.: Shadow prices and well-posedness in the problem of optimal investment and consumption with transaction costs. SIAM J. Control Optim. **51**, 4414–4449 (2013)
7. Collings, P., Haussmann, U.G.: Optimal portfolio selection with transaction costs. In: Chen, S., et al. (eds.) Control of Distributed Parameter and Stochastic Systems, pp. 189–197. Springer, Berlin (1999)
8. Dai, M., Zhong, Y.: Penalty methods for continuous-time portfolio selection with proportional transaction costs. J. Comput. Finance **13**(3), 1–31 (2010)
9. Davis, M.H.A., Norman, A.R.: Portfolio selection with transaction costs. Math. Oper. Res. **15**, 676–713 (1990)
10. Dumas, B., Luciano, E.: An exact solution to a dynamic portfolio choice problem under transactions costs. J. Finance **46**, 577–595 (1991)
11. Evans, J., Henderson, V., Hobson, D.: Optimal timing for an indivisible asset sale. Math. Finance **18**, 545–567 (2008)
12. Guasoni, P., Muhle-Karbe, J.: Portfolio choice with transaction costs: a user's guide. In: Henderson, V., Sircar, R. (eds.) Paris–Princeton Lectures on Mathematical Finance 2013, LNM 2081, pp. 169–201. Springer, Berlin (2013)
13. Herczegh, A., Prokaj, V.: Shadow price in the power utility case. Ann. Appl. Probab. **25**, 2671–2707 (2015)
14. Hobson, D., Tse, A.S.L., Zhu, Y.: Optimal consumption and investment under transaction costs. Math. Finance **29**, 483–506 (2019)
15. Hobson, D., Zhu, Y.: Multi-asset consumption–investment problems with infinite transaction costs. Working paper (2014). Available online at https://arxiv.org/abs/1409.8037
16. Janeček, K., Shreve, S.E.: Asymptotic analysis for optimal investment and consumption with transaction costs. Finance Stoch. **8**, 181–206 (2004)
17. Kallsen, J., Muhle-Karbe, J.: On using shadow prices in portfolio optimization with transaction costs. Ann. Appl. Probab. **20**, 1341–1358 (2010)
18. Liu, H.: Optimal consumption and investment with transaction costs and multiple risky assets. J. Finance **59**, 289–338 (2004)
19. Magill, M.J.P., Constantinides, G.M.: Portfolio selection with transactions costs. J. Econ. Theory **13**, 245–263 (1976)
20. Merton, R.C.: Lifetime portfolio selection under uncertainty: the continuous-time case. Rev. Econ. Stat. **51**, 247–257 (1969)
21. Muthuraman, K., Kumar, S.: Multidimensional portfolio optimization with proportional transaction costs. Math. Finance **16**, 301–335 (2006)
22. Possamaï, D., Soner, H.M., Touzi, N.: Homogenization and asymptotics for small transaction costs: the multidimensional case. Commun. Partial Differ. Equ. **40**, 2005–2046 (2015)
23. Shreve, S.E., Soner, H.M.: Optimal investment and consumption with transaction costs. Ann. Appl. Probab. **4**, 609–692 (1994)
24. Soner, H.M., Touzi, N.: Homogenization and asymptotics for small transaction costs. SIAM J. Control Optim. **51**, 2893–2921 (2013)
25. Tse, A.S.L.: Dynamic economic decision problems under behavioural preferences and market imperfections. PhD Thesis, University of Warwick (2016). Available online at http://wrap.warwick.ac.uk/89266/
26. Whalley, E.A., Wilmott, P.: An asymptotic analysis of an optimal hedging model for option pricing with transaction costs. Math. Finance **7**, 307–324 (1997)